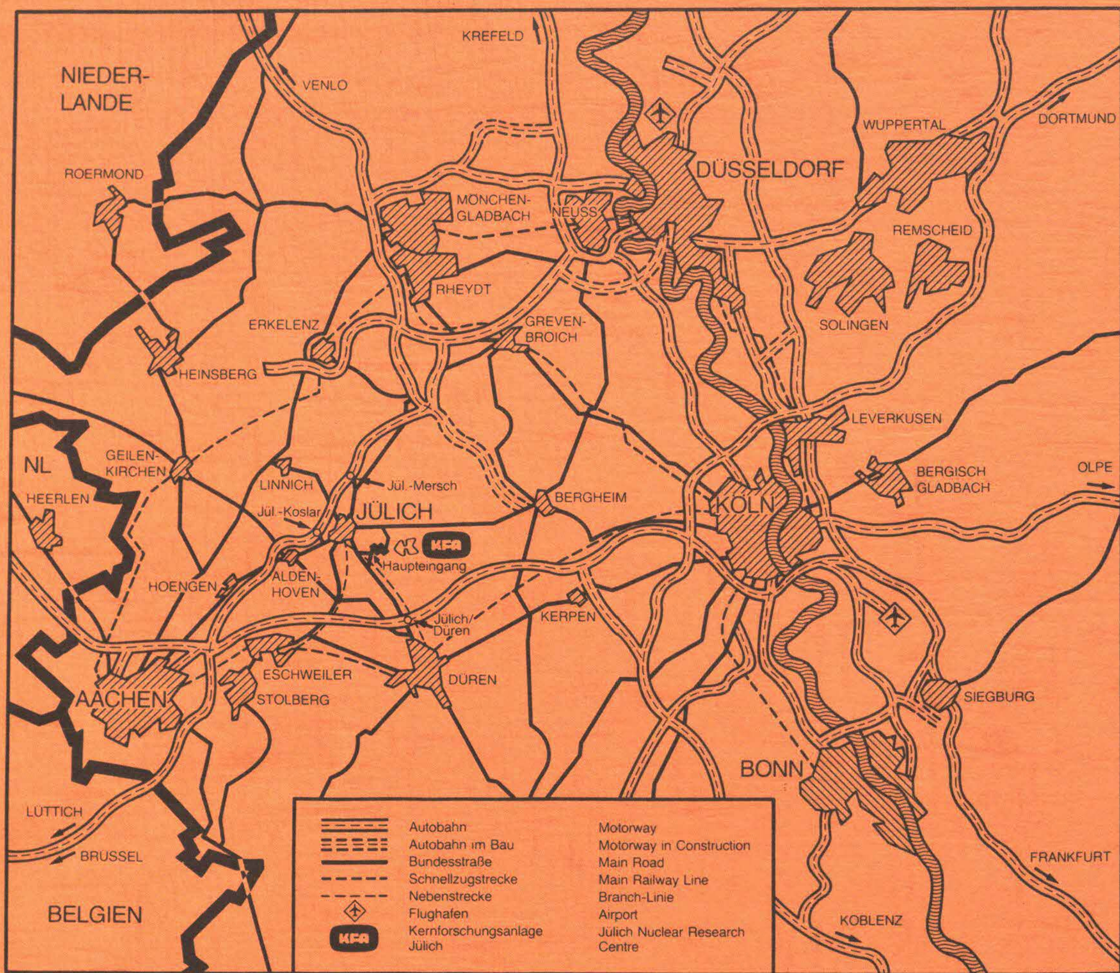
KERNFORSCHUNGSANLAGE JÜLICH GmbH
Institut für Festkörperforschung

Spielstärke und Gewinnchance

von

B. Alefeld

Jül - Spez - 209
Mai 1983
ISSN 0343-7639



Als Manuskript gedruckt

Spezielle Berichte der Kernforschungsanlage Jülich – Nr. 209
Institut für Festkörperforschung Jül - Spez - 209

Zu beziehen durch: ZENTRALBIBLIOTHEK der Kernforschungsanlage Jülich GmbH
Postfach 1913 · D-5170 Jülich (Bundesrepublik Deutschland)
Telefon: 02461/610 · Telex: 833556 kfd

Spielstärke und Gewinnchance

von

B. Alefeld

SPIELSTÄRKE UND GEWINNCHANCE

B. Alefeld

Institut für Festkörperforschung
KFA, D-5170 Jülich, Fed. Rep. Germany

INHALT:

Einführung

Die mathematische Statistik hat in der Physik zur Beschreibung vieler Vorgänge große Bedeutung erlangt. In dieser Arbeit wird diese Methode zur Untersuchung der Gewinnchance bei Sportspielen angewandt. Als spezielles Beispiel untersuchen wir die Sportart Tennis, die sich für statistische Untersuchungen besonders gut eignet, da bei einem Tennismatch relativ große Zahlen auftreten. Man wird zunächst einwenden, daß solche Überlegungen nur wenig Praxisbezug haben, da Tennis kein Würfelspiel ist, sondern der Mensch die zentrale Rolle spielt. Dies ist sicher richtig, doch entgegenhalten möchte ich, daß bei sehr komplexen wissenschaftlichen Problemen wie der Psychologie oder der Verhaltensforschung, die eine kausale Beschreibung in Form mathematischer Gleichungen nicht zulassen, statistische Untersuchungen von Zusammenhängen (Korrelationen) eine sehr große Bedeutung erlangt haben. Tennis ist ebenfalls ein sehr komplizierter Vorgang, es gibt aber auch hier einige bisher nicht oder nur intuitiv verstandene Aspekte, die einer mathematisch statistischen Beschreibung zugänglich sind und auch experimentell geprüft werden können.

1. Die Tenniszählweise

Die Regeln des Tennisspiels sind für die folgenden Überlegungen unerheblich, wesentlich ist nur, daß von zwei Spielern, die im folgenden mit A und B bezeichnet werden, Punkte erspielt werden und nach einem speziellen Schema zum Endresultat aufaddiert werden. Punkte werden zu Spielen addiert, Spiele zu Sätzen und die Anzahl der Gewinnsätze (zwei oder drei) entscheidet letztlich über den Ausgang des Matches. Dieses Additionsschema ist keinesfalls logisch zwingend, denn in vergleichbaren Sportarten wie Tischtennis oder Spuash hat man sich andere Regeln für die Addition der Punkte gegeben, und in Sportarten wie Handball, Fußball oder Eishockey gewinnt derjenige, der in einer vorgegebenen Zeit die meisten "Punkte" (Tore)erzielt. Die statistischen Folgerungen aus der Zählweise einer Sportart können recht unterschiedlich sein. Dies allgemein zu zeigen wäre sicherlich interessant, würde hier aber viel zu weit führen. Wir beschränken uns im folgenden ausschließlich auf Tennis.

1.1 Wahrscheinlichkeiten für den Punkt-, Spiel-, Satz- und Matchgewinn

Wir machen in diesem Abschnitt die einzige sehr wichtige Annahme, daß Spieler A während des gesamten Matches einen Punkt mit der Wahrscheinlichkeit p (point) gewinnt, wobei p zwischen 0 und 1 liegt. Es folgt sofort, daß sein Gegenspieler B jeden Punkt mit der Wahrscheinlichkeit 1-p gewinnt. Betrachten wir z.B. den konkreten Wert p = 0,55, so heißt das in der Praxis, daß von 20 ausgespielten Punkten Spieler A im Mittel 11 Punkte gewinnt und Spieler B neun Punkte gewinnt. Wir fragen nun, mit welcher Wahrscheinlichkeit Spieler A ein Spiel, einen Satz oder das gesamte Match gewinnt. Fig. 1 zeigt das Flußdiagramm eines Spiels. Der Spielstand nach Ausspielung des ersten Punktes ist 15:0 für Spieler A mit der Wahrscheinlichkeit p, oder 0:15 für Spieler B mit der Wahrscheinlichkeit 1-p. Nach dem zweiten Punkt gibt es drei mögliche Spielstände, nämlich 30:0, 15:15 oder 0:30. Die Wahrscheinlichkeit, daß Spieler A den Spielstand 30:0 erreicht, ist das Produkt der Wahrscheinlichkeiten den ersten sowie den zweiten Punkt zu erhalten, d.h. $W_{30:0} = p^2$ ($W_{30:0}$: Wahrscheinlichkeit für Spieler A, nach zwei Punkten den Spielstand 30:0 zu erreichen). Ebenso erhält man für Spieler B, den Spielstand 0:30 zu erreichen, zu $W_{0:30} = (1-p)^2$. Wichtig ist nun, daß der Spielstand 15:15 auf zwei verschiedenen Wegen erreicht werden kann, nämlich über 15:0 nach 15:15 oder über 0:15 nach 15:15. Die Einzelwahrscheinlichkeiten, diesen Spielstand über die verschiedenen Wege zu erreichen, sind gleich, nämlich $p(1-p)$, die Gesamtwahrscheinlichkeit für diesen Spielstand ist die Summe der Einzelwahrscheinlichkeiten, also $2\,p(1-p)$. Am Spielstand 40:30 sei nochmals demonstriert, wie man die Gesamtwahrscheinlichkeit für das Erreichen dieses Spielstands berechnet. Die Einzelwahrscheinlichkeit für jeden der möglichen Wege ist $p^3(1-p)^2$, wobei die beiden Exponenten angeben, daß man in Fig. 1 dreimal nach links unten (Punktgewinn für A) und zweimal nach rechts unten (Punktgewinn für B) gehen muß. Es gibt, wie man leicht nachprüfen kann, zehn verschiedene Wege zum Spielstand 40:30 zu kommen, so daß man für $W_{40:30}$ den Wert 10 $p^3(1-p)^2$ erhält. Die Wahrscheinlichkeiten für das Erreichen eines bestimmten Spielstands sind in Fig. 1 jeweils unter dem entsprechenden Spielstand angegeben. Beim Erreichen des Spielstands G (Game) wird dieses Schema abgebrochen. Nach unten setzt sich das Flußdiagramm endlos fort Um nun die Wahrscheinlichkeit g zu erhalten, mit der Spieler A das Spiel gewinnt, müssen alle Einzelwahrscheinlichkeiten dafür, das Spiel auf irgendeine Weise zu gewinnen, addiert werden. Man erhält also

$$g = p^4+4p^4(1-p)+10p^4(1-p)^2+20p^5(1-p)^3+40p^6(1-p)^4 + 80p^7(1-p)^5+160p^8(1-p)^6+320p^9(1-p)^7 \;.... \qquad (1)$$

$$= p^4+4p^4(1-p)+10p^4(1-p)^2 \left[1+2p(1-p)+4p^2(1-p)^2...\right] \qquad (2)$$

Der Ausdruck in der eckigen Klammer ist eine geometrische Reihe mit der Summe $[....] = 1/\left(1-2p(1-p)\right)$. Damit erhält man für g

$$g = p^4\left(1 + 4(1-p) + \frac{10(1-p)^2}{1-2p(1-p)}\right) \qquad (3)$$

In Fig. 2 wurde g (multipliziert mit 100, d.h. in Prozent) als Funktion von p aufgetragen. Wir können z.B. für p = 55 den Wert 62 ablesen, d.h., wenn Spieler A im statistischen Mittel von 100 Punkten 55 Punkte gewinnt, so gewinnt er von 100 Spielen im statistischen Mittel 62 Spiele. Die Überlegenheit von Spieler A bezüglich eines Spielgewinns ist größer als die Überlegenheit bezüglich eines Punktgewinns. Diese Tatsache bezeichnen wir als Verstärkungseffekt.

Die nächste Frage lautet nun, mit welcher Wahrscheinlichkeit gewinnt A einen Satz. Fig. 3 zeigt das Flußdiagramm eines Satzes. Nur an dieser Stelle geht unsere Näherung ein, d.h.,

Fig. 3 und die daraus abgeleitete Gleichung (4) sind nur dann richtig, wenn der Aufschlagwechsel und der Wechsel der Platzseiten keine Rolle spielt (Aufschlageinfluß siehe Abschnitt 1.3). Für die Wahrscheinlickeit, daß Spieler A ein Spiel gewinnt, haben wir also g einzusetzen, dafür, daß er ein Spiel verliert, setzen wir 1-g ein. Außerdem wird beim Stande von 6:6 Tie-Break gespielt. Es sei dem Leser überlassen zu zeigen, daß Spieler A für $p > 0{,}5$ ein Tie-Break-Spiel mit einer etwas größeren Wahrscheinlichkeit gewinnt als ein normales Spiel. Wir lassen dies im weiteren außer acht, da dies für die späteren Betrachtungen sicherlich unerheblich ist. Addiert man nun wiederum die Wahrscheinlichkeiten, daß Spieler A den Satz überhaupt gewinnt und bezeichnen wir diese Wahrscheinlichkeit mit s (set), so erhalten wir

$$s = g^6+6g^6(1-g)+21g^6(1-g)^2+56g^6(1-g)^3+126g^6(1-g)^4 +252g^7(1-g)^5+504g^7(1-g)^6. \qquad (4)$$

s ist (mit 100 multipliziert) wiederum in Fig. 2 abhängig von p aufgetragen und wir können für p = 55 den Wert 81 ablesen, d.h., Spieler A gewinnt von 100 Sätzen 81, d.h., auch hier tritt ein großer Verstärkungseffekt auf.

Fig. 4 zeigt das Flußdiagramm für ein Match; wir haben hier für die Wahrscheinlichkeit, daß A den Satz gewinnt bzw. verliert die Werte s bzw. 1-s einzusetzen. Für die Wahrscheinlichkeit m_2, das Match mit 2 Gewinnsätzen zu gewinnen bzw. m_3, das Match mit 3 Gewinnsätzen zu gewinnen, erhalten wir also

$$m_2=s^2+2s^2(1-s) \qquad (5)$$

und

$$m_3=s^3+3s^3(1-s)+6s^3(1-s)^2. \qquad (6)$$

In Fig. 2 sind auch diese Größen abhängig von p aufgetragen und wir erhalten für den Wert p =   55, daß Spieler A mit 91 % Wahrscheinlichkeit ein Match mit 2 Gewinnsätzen und mit 95 % Wahrscheinlichkeit ein Match mit 3 Gewinnsätzen gewinnt. Wir stellen also fest, daß eine kleine Überlegenheit beim Ausspielen eines Punktes (z.B. p = 0,55) zu einer großen Überlegenheit bezüglich des Matchgewinns (m_2=0,91; m_3=0,95) führt.

1.2 Der Verstärkungseffekt als Ursache der Fehleinschätzung

Wir definieren nun den Verstärkungsfaktor G als das Verhältnis der gewonnenen zu den verlorenen Begegnungen bezogen auf das Verhältnis der gewonnenen zu den verlorenen Punkten, also

$$G = \frac{m}{1-m} : \frac{p}{1-p} \tag{7}$$

In Fig. 5 ist G als Funktion von p aufgetragen. Wie ungeheuer wichtig dieser Verstärkungsfaktor ist, soll an folgendem Beispiel nochmals verdeutlicht werden. Spieler A möge von 40 Punkten 21 gewinnen und 19 verlieren, was einem Wert p = 0,525 entspricht. Setzen wir diesen Wert und den entsprechenden Wert m in Gleichung (7), so erhalten wir für ein Match mit drei Gewinnsätzen den Verstärkungsfaktor G = 3,55, d.h. mit anderen Worten, wenn man von 40 Punkten 21 Punkte gewinnt, ist das Verhältnis der gewonnenen zu den verlorenen Begegnungen gleich

$$\frac{m}{1-m} = G\,\frac{p}{1-p} = 3,92 \approx 4$$

Das heißt, von 5 Begegnungen gewinnt A 4 Begegnungen.
Es wird später gezeigt, daß diese Aussage in guter Näherung auch bei Berücksichtigung des Aufschlageinflusses gilt.
Der überraschend große Verstärkungseffekt vom Punktgewinn zum Matchgewinn ist sicherlich die Ursache für die häufig

auftretende Fehleinschätzung von Tennisspielern. Man kann oft beobachten, daß Spieler B behauptet, er spiele eigentlich gar nicht schlechter als A, und trotzdem verliere er doch meistens. Der Grund für diese Fehleinschätzung liegt darin, daß es sehr schwer ist festzustellen, ob man von 40 Punkten 20 oder nur 19 Punkte gewinnt. Dieser kleine Unterschied kommt erst im Laufe der Begegnung zum Tragen und führt dazu, daß man mit 21 Gewinnpunkten in 4 von 5 Begegnungen Sieger bleibt, mit 20 Gewinnpunkten natürlich die Hälfte aller Begegnungen gewinnt und mit 19 Gewinnpunkten bei 4 von 5 Begegnungen verliert.

Die Tatsache, daß nahezu unsichtbare Unterschiede im Punktgewinn zu großen Unterschieden im Matchgewinn führen, ist vermutlich auch die Erklärung für folgendes Phänomen. Spieler A gewinnt leicht gegen B, B gewinnt leicht gegen C und C gewinnt leicht gegen A. Wenn die drei Spieler in der Reihenfolge A - B - C in einer Rangliste stehen, so wird C durch B abgeblockt (C glaubt, er müßte zwei Plätze höher stehen). Die richtige Aussage ist aber, daß die drei Spieler gleich gut sind. Daraus folgt aber nicht, daß jeder gegen jeden gleich häufig gewinnt. Geringfügige Unterschiede der individuellen Spielweise oder auch psychologische Gründe können das Gleichgewicht in den Einzelbegegnungen stark zur einen oder anderen Seite verschieben.

1.3 Berücksichtigung des Aufschlageinflusses

Wie schon erwähnt, ist die Annahme, daß Spieler A während des gesamten Matches einen Punkt mit der Wahrscheinlichkeit p gewinnt und Spieler B einen Punkt mit der Wahrscheinlichkeit 1-p gewinnt in den höheren Spielklassen sicher unrealistisch. Wir werden deshalb im folgenden die Annahme machen, daß Spieler A nur während seines Aufschlags den Punkt mit der Wahrscheinlichkeit p gewinnt, während Spieler B bei eigenem

Aufschlag den Punkt mit der Wahrscheinlichkeit q gewinnt. Die Spiele werden entsprechend mit den Wahrscheinlichkeiten g und h gewonnen. Das entsprechende Flußdiagramm für den Ablauf eines Satzes ist in Fig. 6 angedeutet. Spieler A beginnt mit dem Aufschlag und gewinnt das 1. Spiel mit der Wahrscheinlichkeit g. Anschließend schlägt Spieler B auf und gewinnt das zweite Spiel mit der Wahrscheinlichkeit h. Man kann nun für alle Kombinationen der möglichen Werte von g und h die Wahrscheinlichkeiten berechnen, daß Spieler A den Satz gewinnt. Die mathematische Berechnung dieser Wahrscheinlichkeiten ist nicht schwerer als in dem einfachen Fall des ersten Abschnitts, aber doch wesentlich umfangreiche Deshalb zeigen wir hier nur die Ergebnisse. In Fig. 7 ist auf der horizontalen Achse die Größe g und auf der vertikalen Achse die Größe h aufgetragen. Ein spezielles Wertepaar von g und h entspricht einem Punkt in dieser(g,h)-Ebene. Liegt dieser Punkt z.B. auf der Kurve, die mit 0,6 gekennzeichnet ist, so gewinnt Spieler A den Satz gerade mit dieser Wahrscheinlichkeit 0,6. Das heißt, die Kurven sind Kurven gleicher Wahrscheinlichkeit für Spieler A, den Satz zu gewinnen. Es ist klar, daß für g = h jeder Spieler den Satz mit 50 % Wahrscheinlichkeit gewinnt. Dies zeigt die Gerade, die mit 0,5 gekennzeichnet ist. Der Spezialfall des 1. Abschnitts muß natürlich in diesen Kurven enthalten sein, er besagt, daß für einen gegebenen Wert g für h der Wert 1-g gewählt werden muß. Dieser Fall ist als gestrichelte Gerade eingezeichnet und ein Vergleich mit Fig.2 zeigt Übereinstimmung. In Fig. 7 wurden noch 3 Bereiche eingezeichnet, die mit 1, 2 und 3 gekennzeichnet sind. Im Bereich 1 ist der Aufschlag ein Nachteil, es ist gewissermaßen der Anfängerbereich. Im Bereich 2 ist der Aufschlag weder ein Vor- noch ein Nachteil, dies ist der Bereich mittelmäßiger Tennisspieler. Im Bereich 3 haben beide Spieler einen großen Aufschlagvorteil, hier sind die sehr guten Spieler zu suchen.

Wir zeigen nun, in welcher Weise der Verstärkungsfaktor berechnet werden muß, wenn Spieler A sein Aufschlagspiel mit der Wahrscheinlichkeit g gewinnt und Spieler B entsprechend mit h gewinnt. Aus Fig. 7 können wir entnehmen, daß in dem Bereich $0{,}3 < g,\ h < 0{,}7$ in guter Näherung gilt:

$$s(g',1-g') \approx s\big(g'+a, 1-(g'+a)\big)$$

Mit $g = g'+a$ und $h = 1-(g'+a)$ erhalten wir durch Elimination von a für den Wert g'

$$g' = 0{,}5 + \frac{g-h}{2} \text{ und entsprechend } 1-g' = 0{,}5 - \frac{g-h}{2}$$

Gewinnt Spieler A sein Aufschlagspiel mit g und B entsprechen mit h, so kann man s aus der einfachen Gl. (4) berechnen, indem man g' einsetzt. Das heißt mit anderen Worten, bei der Berechnung von s geht nur die Differenz des Aufschlagvorteils ein, das Problem kann also so behandelt werden, als ob A alle Spiele mit der gleichen Wahrscheinlichkeit $g' = 0{,}5 + \frac{g-h}{2}$ gewinnen würde. Die Berechnung des Verstärkungsfaktors G erfolgt dann, wie in 1.2 gezeigt wurde.

2. Häufigkeits- und Wahrscheinlichkeitsverteilungen

2.1 Punkthäufigkeiten und die Qualität des Aufschlags

Ein Tennisspieler möge in einem Match n-mal aufschlagen und dabei n_+ Gewinnpunkte und n_- Verlustpunkte erzielen. n_+ und n_- werden als absolute Häufigkeiten der Gewinn- bzw. Verlustpunkte bezeichnet. Das Verhältnis $n_+ : n$ wird als relative Häufigkeit der Pluspunkte bezeichnet. Es ist äquivalent der Wahrscheinlichkeit, bei einem Aufschlag einen Punkt zu gewinnen. Für die Teilnehmer an den Endspielen der Herren der Wimbledon-Turniere von 1981 und 1982 sind die Werte von n_+, n_-, p (erstgenannter Spieler) und q (zweitgenannter Spieler) und die aus Gl. (3) berechneten Werte für g und h angegeben.

Jahr	Teil-nehmer	n_+	n_-	p, q $\frac{n_+}{n_+ + n_-}$	g, h
1981	ME	108	56	$0{,}66 \pm 0{,}06$	$0{,}85^{+0{,}08}_{-0{,}11}$
	B	99	53	$0{,}65 \pm 0{,}06$	$0{,}83^{+0{,}08}_{-0{,}11}$
1982	ME	107	53	$0{,}67 \pm 0{,}067$	$0{,}86^{+0{,}08}_{-0{,}12}$
	C	118	68	$0{,}63 \pm 0{,}06$	$0{,}80^{+0{,}09}_{-0{,}11}$

Tab. 1: Die Bestimmung der Werte p, q, g und h der Endspielteilnehmer von Wimbledon im Herreneinzel von 1981 u. 1982.

n_+ (n_-): Anzahl der gewonnenen (verlorenen) Punkte bei eigenem Aufschlag. Tie-Break berücksichtigt.
ME: McEnroe B: Borg C: Connors.

Wir sehen zunächst, daß die Werte für p und q sowie g und h nahezu gleich sind. Außerdem sehen wir, daß die statistische Unsicherheit der Werte aufgrund der kleinen Datenmenge doch beträchtlich ist.

Tab. 1 bringt also zum Ausdruck, daß die beteiligten Spieler bei eigenem Aufschlag im Mittel mit 65 % einen Punkt gewinnen und mit etwa 83 % ihr Aufschlagspiel gewinnen.

2.2 Spielverteilung P_i^G und P_i^H

Ein Tennisspiel kann mit 4, 5, 6, 8, 10 ... i ... Punkten gespielt werden. Die Größe P_i^G gibt an, mit welcher Wahrscheinlichkeit ein Spiel, in dem ein Spieler mit der Wahrscheinlichkeit p einen Punkt gewinnt, mit i Punkten endet. Aus Fig. 1 können wir direkt ablesen:

$$\begin{aligned}
P_4^G &= p^4 + (1-p)^4 \\
P_5^G &= 4\left((p^4(1-p) + p\,(1-p)^4)\right) \\
P_6^G &= 10\left((p^4(1-p)^2 + p^2\,(1-p)^4)\right) \\
P_8^G &= 20\left((p^5(1-p)^3 + p^3\,(1-p)^5)\right)
\end{aligned} \tag{8a}$$

$$P^G_{2j-2} = 5 \cdot 2^{j-3}\left((p^j(1-p)^{j-2} + p^{j-2}.(1-p)^j)\right)$$

Entsprechend gilt:

$$P^H_4 = q^4 + (1-q)^4 \quad \text{usw.} \qquad (8b)$$

	n_4	n_5	n_6	n_8	n_{10}	n_{12}	n_{14}	n_{16}	n_{18}	n_{20}	Σ
1981	5	13	13	6	3	2	1	0	1	0	44
1982	8	19	15	4	1	3	1	1	0	0	52
81/82	13	32	28	10	4	5	2	1	1	0	96
P^G_i (81/82)	13.5	33.3	29.2	10.4	4.2	5.2	2.1	1.0	1.0	0	100
P^G_i (ber.)	19,7	29	28	12.7	5.8	2.6	1.2	0.5	0.2	0	100

Tabelle 2
Die absoluten Häufigkeiten n_i der Spiele, die mit i Punkten gespielt wurden, von den Endspielen der Herren der Wimbledon-Turniere von 1981 und 1982. P^G_i ist die relative Häufigkeit der Spiele mit i Punkten.

In Tab. 2 wurden sowohl die Werte der Endspiele von Wimbledon 1981 und 1982 eingetragen als auch die aus der Gl. (8) mit dem Wert p = 0,65 berechneten Werte angegeben. Fig. 8 zeigt eine graphische Darstellung der Resultate (die Wimbledon-Endspiele wurden in einem Punkt zusammengefaßt). Die durchgezogene Kurve ist nur für die Augenführung eingezeichnet, sie verbindet die berechneten Häufigkeiten. Wir sehen, daß die 4 : 0-Ergebnisse zu selten und die 4 : 1-Ergebnisse etwas häufiger auftreten als wir berechnet haben. Sonst ist die Übereinstimmung sehr gut. Die Abweichung bei den 4 : 0-Ergebnissen ist so groß, daß Zweifel angebracht sind, ob sie rein statistischer Natur ist. Wäre die Abweichung rein statistischer Natur, so sollte man bei 93 von 100 ähnlichen Turnieren eine größere Häufigkeit der 4 : 0 Ergebnisse erwarten.

Wir sind also geneigt anzunehmen, daß diese Abweichungen der 4 : 0- und 4 : 1-Häufigkeiten von den erwarteten Häufigkeiten der erste Hinweis auf einen psychologischen Effekt dergestalt ist, daß ein Spieler, der mit 3 : 0-Punkten führt, eher einmal einen

Fehler riskiert als dies im Mittel der Fall ist, oder, daß der 1. Punkt eines Spieles vielmehr umkämpft ist als die weiteren Punkte. Wir gehen hier nicht weiter auf dieses Problem ein, wir stellen nur fest, daß mit mehr statistischem Material vielleicht schon an dieser Stelle gewisse Feinheiten des Tennisspiels, die man sonst nie sehen würde, aufgedeckt werden könnten.

2.3 Satzhäufigkeiten

Ein Satz kann mit 6, 7, 8, 9, 10, 12 Spielen oder mit Tie-Break enden. Wir bezeichnen die Wahrscheinlichkeiten für diese Satzausgänge mit P_i^S; der Satz mit Tie-Break bekommt den Index i = 13 Wir nehmen wieder an, daß Spieler A seine Aufschlagspiele mit der Wahrscheinlichkeit g und Spieler B seine Aufschlagspiele mit der Wahrscheinlichkeit h gewinnt. Wenn wir das allgemeine Flußdiagramm der Fig. 6 explizit ausrechnen, so können wir folgende vier Wahrscheinlichkeiten unterscheiden:

$$ {}^{A}_{A}P_i^S \qquad {}^{A}_{B}P_i^S \qquad {}^{B}_{B}P_i^S \quad \text{und} \quad {}^{B}_{A}P_i^S $$

Der linke obere Index gibt an, welcher Spieler im 1. Spiel des Satzes Aufschlag hat, der linke untere Index unterscheidet zwischen den Gewinnsätzen der Spieler A und B. In Tab. 3 sind die Werte dieser Größen für die drei verschiedenen Wertepaare g = h = 0,5; g = h = 0,8 und g = 0,8, h = 0,6 eingetragen.

Wir betrachten die drei Beispiele zunächst getrennt:

1. g = h = 0,5
   Die Satzverteilung ist für beide Spieler gleich und unabhängig davon, wer mit dem Aufschlag beginnt. Die 5. Zeile zeigt, daß die häufigsten Sätze mit 6 : 4 enden, die 7 : 5- und Tie-Break-Ergebnisse sind gleichwahrscheinlich. Bemerkenswert ist, daß das Ergebnis 6 : 2 häufiger auftritt als das Ergebnis 7 : 5 oder Tie-Break. Die letzte Spalte zeigt, daß jeder Spieler den Satz mit 50 % Wahrscheinlichkeit gewinnt.

Zeile Nr.	g und h	Index A u. B	i							Σ
			6	7	8	9	10	12	13	i
1		A, A	1,6	4,7	8,2	10,9	12,3	6,2	6,2	50
2		A, B	1,6	4,7	8,2	10,9	12,3	6,2	6,2	50
3	0,5	B, B	1,6	4,7	8,2	10,9	12,3	6,2	6,2	50
4	0,5	B, A	1,6	4,7	8,2	10,9	12,3	6,2	6,2	50
5		$\frac{1}{2}\Sigma$	3,1	9,4	16,4	21,9	24,6	12,3	12,3	100
6		A, A	0,4	4,2	4,0	18,8	6,7	5,1	10,9	50
7	0,8	A, B	0,4	1,0	7,1	4,7	20,8	5,1	10,9	50
8		B, B	0,4	4,2	4,0	18,8	6,7	5,1	10,9	50
9	0,8	B, A	0,4	1,0	7,1	4,7	20,8	5,1	10,9	50
10		$\frac{1}{2}\Sigma$	0,8	5,2	11,0	23,5	27,5	10,2	21,7	100
11		A, A	3,3	13,8	12,2	24,6	10,5	6,9	9,7	80,9
12		A, B	0,2	0,5	2,7	2,3	8,6	2,6	2,3	19,1
13	0,8	B, B	0,2	1,5	1,7	6,9	3,9	2,6	2,3	19,1
14	0,6	B, A	3,3	6,9	19,1	12,3	22,8	6,9	9,7	80,9
15		$\frac{1}{2}\Sigma$	3,4	11,3	17,8	23,1	22,9	9,4	12	100
16 Wimbledon 77-82			2,3	8,4	17	22,9	24,7	8,9	15,8	
17 $0{,}35\cdot\Sigma 10+0{,}65\cdot\Sigma 15$			2,5	9,16	15,4	23,24	24,5	9,6	15,4	

Tabelle 3 Satzhäufigkeiten
Wir erklären diese Tabelle an einem konkreten Beispiel. Wir betrachten z.B. die Zeile 12. Spieler A gewinnt sein eigenes Aufschlagspiel mit 80 % Wahrscheinlichkeit, während Spieler B sein Aufschlagspiel mit 60 % Wahrscheinlichkeit gewinnt. Speiler A beginnt mit dem Aufschlag, Spieler B gewinnt 0,2 % der Sätze mit 6 : 0, 0,5 % der Sätze mit 6 : 1, 2,7 % der Sätze mit 6 : 2 usw. Die letzte Spalte gibt an, daß Spieler B mit 19,1 % Wahrscheinlichkeit den Satz gewinnt. Die 15. Zeile gibt die Satzhäufigkeiten an, wenn man über A und B ermittelt. Die Zeile 16 gibt die relativen Satzhäufigkeiten aus den ungesetzten Begegnungen der Herren aus der 1. Runde der Wimbledon-Turniere von 1977 bis 1982 an. Zeile 17 gibt eine spezielle Satzhäufigkeitsverteilung an, die man dann erhält, wenn man annimmt, daß 35 % der Begegnungen durch Zeile 10 und 65 % der Begegnungen durch Zeile 15 der Tabelle beschrieben werden können. Die Werte der Zeilen 16 und 17 sind in Fig. 9 aufgetragen.

2. g = h = 0,8
Aus der letzten Spalte entnehmen wir, daß auch hier jeder Spieler, so wie es sein muß, den Satz mit 50 % Wahrscheinlichkeit gewinnt. Die Satzverteilung hängt jedoch davon ab, wer mit dem Aufschlag beginnt. Beginnt z.B. Spieler A mit dem Aufschlag, so gewinnt dieser 18,8 % der Sätze mit 6 : 3 und 6,7 % der Sätze mit 6 : 4. Beginnt dagegen B mit dem Aufschlag, so gewinnt Spieler A nur 4,7% der Sätze mit 6 : 3, aber 20,8 % mit 6 : 4.

3. g = 0,8, h = 0,6
Aus der letzten Spalte entnehmen wir, daß die Wahrscheinlichkeit, den Satz zu gewinnen, unabhängig davon ist, wer mit dem Aufschlag beginnt, Spieler A gewinnt mit 80,9 % Spieler B mit 19,1 %. Außerdem sieht man, daß im Gegensatz zum 1. Fall hier alle vier Unterverteilungen (Zeile 11 bis 14) voneinander verschieden sind.

Wir betrachten noch die über den Aufschlag gemittelten Häufigkeitsverteilungen (Zeile 5, 10 und 15). Für g = h = 0,8 ist die Häufigkeit der Tie-Break-Sätze wesentlich größer und die Zahl der 6 : 0- und 6 : 1-Sätze wesentlich kleiner als im Falle g = h = 0,5. Für g = 0,8 und h = 0,6 (unausgeglichene Paarung) steigt die Anzahl der 6 : 0- und 6 : 1-Ergebnisse an.

Um nun zu sehen, wie realistisch unsere Rechnungen sind, wurden alle Satzhäufigkeiten der ungesetzten Paarungen aus der 1. Runde der Wimbledon-Turniere von 1977 bis 1982 bestimmt und in Zeile 16 der Tab. 3 eingetragen. Vergleichen wir diese Werte mit den oben berechneten Wahrscheinlichkeiten der Zeilen 5, 10 und 15, so sehen wir in jedem Fall deutliche Abweichungen. Nehmen wir hingegen den speziellen Wert (0,35 · Zeile 10 + 0,65 · Zeile 15), der in Zeile 17 eingetragen wurde, so wird die Überein-stimmung zwischen den Wimbledon-Resultaten und den berechneten Häufigkeiten sehr gut. Dies zeigt nochmals Fig. 9.

Natürlich ist unsere Mittelung über nur zwei verschiedene Spielstärkepaarungen nicht exakt; wir glauben aber ganz sicher, daß eine exakte Mittelung mit einer noch zu bestimmenden Spielstärkeverteilung H(g,h), also

$$\bar{P}_i^S = \frac{1}{2} \int_0^1 H(g,h) \left({}^{A}_{B}P_i^S + {}^{A}_{B}P_i^S + {}^{B}_{B}P_i^S + {}^{B}_{A}P_i^S \right) dg\,dh \qquad (9)$$

die Wimbledon-Resultate sehr genau beschreibt.

Da wir keine genaueren Kenntnisse von H(g,h) haben, verzichten wir hier auf eine weitere Untersuchung dieser Mittelung.

2.4 Die Häufigkeitsverteilung der Matchergebnisse

Ein Tennismatch mit drei Gewinnsätzen kann entweder mit 3, 4 oder 5 Sätzen gespielt werden. Wir nehmen an, daß Spieler A einen Satz mit der Wahrscheinlichkeit s und entsprechend Spieler B mit der Wahrscheinlichkeit 1-s gewinnt. (Im vorhergehenden Abschnitt hatten wir für eine Spielerpaarung mit g = 0,8 und h = 0,6 den Wert s = 0,809 erhalten.) Aus Fig. 4 erhalten wir für die Häufigkeiten der drei oben genannten Matchausgänge

$$\begin{aligned} P_3^M &= s^3 + (1-s)^3 \\ P_4^M &= 3\left(s^3(1-s) + s(1-s)^3 \right) \qquad (10) \\ P_5^M &= 6\left(s^3(1-s)^2 + s^2(1-s)^3 \right) \end{aligned}$$

In Fig. 10 sind diese drei Häufigkeiten in Abhängigkeit von s aufgetragen. (Es gilt: $P_i^M(s) = P_i^M(1-s)$, deshalb wurde nur der Bereich $0{,}5 \leq s \leq 1$ angegeben.) Wir sehen also, daß für s = 0,5 (d.h. bei gleich guten Spielern)

25 % aller Begegnungen mit 3 Sätzen und jeweils 37,5 % der Begegnungen mit 4 bzw. 5 Sätzen enden. Mit wachsendem Spielstärkeunterschied nehmen die 3 : 2-Ergebnisse schneller ab als die 3 : 1-Ergebnisse, die 3 : 0-Ergebnisse nehmen natürlich in demselben Maße zu wie die beiden anderen Ergebnisse abnehmen.

Tab. 4 enthält die erzielten absoluten Häufigkeiten aller ungesetzten Begegnungen aus der 1. Runde der Wimbledon-Turniere von 1977 bis 1982. Der Mittelwert der relativen Häufigkeiten ist in der letzten Spalte angegeben. Vergleichen wir nun diese gemittelten Häufigkeiten mit unserer Rechnung, so entnehmen wir aus Fig. 10, wo die Wimbledonresultate als Punkte eingezeichnet sind, daß die Übereinstimmung für s = 0,76 ganz gut ist. Es ist zunächst sehr unwahrscheinlich, daß alle Begegnungen in Wimbledon mit einem einzigen Wert s = 0,76 beschrieben werden können. Es ist anzunehmen, daß eine Spielstärkeverteilung H(s) vorliegt, wo s zwischen 0,5 (völlig ausgeglichen) und 1 (völlig unausgeglichen) liegt.

Durch Mittelung erhalten wir

$$\overline{P_i^M} = \int_{0,5}^{1} H(s)\, P_i^M(s)\, ds \qquad (11)$$

$$\text{mit} \int_{0,5}^{1} H(s)\, ds = 1$$

Wir behandeln zunächst den Fall, daß die Spielerpaarungen in dem Bereich s = 0,5 und s_{max} gleich häufig vorkommen, also

$$H(s) = \frac{1}{s_{max}-0,5} \quad \text{für} \quad 0,5 \leq s \leq s_{max}$$

$$= 0 \quad \text{für} \quad s_{max} < s \leq 1 \qquad (11\ a)$$

	77	78	79	80	81	82	P_i^M (%)
n_3^M	22	22	26	20	19	23	45,4
n_4^M	15	15	15	19	22	13	35
n_5^M	11	11	6	11	8	10	19,6

Tab. 4

Absolute Häufigkeiten n_i^M und relative, gemittelte Häufigkeiten P_i^M der Drei-, Vier- und Fünfsatzergebnisse der ungesetzten Begegnungen der 1. Runde der Wimbledon-Turniere von 1977 bis 1982.

In Fig. 11 sind die so ermittelten Häufigkeiten als Funktion von s_{max} aufgetragen. Die beste Übereinstimmung mit den Wimbledon-Resultaten, die als Punkte eingetragen sind, findet man für den Wert $s_{max} = 0{,}95$. Es ist sehr bemerkenswert, daß die Annahme einer einigermaßen ausgeglichenen Spielstärkeverteilung ($s \approx 0{,}5$) weit von der Wahrheit entfernt ist. Es muß sogar angenommen werden, daß ein großer Teil aller Spielpaarungen sehr unausgeglichen ist.

Man könnte natürlich auch eine Spielstärkeverteilung untersuchen, die bei einem mittleren Wert $\bar{s}$ ein Maximum hat und nach beiden Seiten abfällt. In diesem Fall hätten wir eine zweiparametrige Spielstärkeverteilung (Lage und Breite), und es ist völlig klar, daß auch damit die Wimbledon-Resultate sehr gut erklärt werden können. Wir vermuten, daß diese letzte Möglichkeit der Wahrheit am nächsten kommt. Wenn wir nämlich genauer untersuchen, wie groß der Abstand der zusammengelosten Spieler in der ATP-Rangliste ist, so stellen wir fest, daß man bei den meisten Paarungen von einem eindeutig besseren und einem eindeutig schlechteren Spieler und nur in seltenen Fälle von zwei nahezu gleichguten Spielern sprechen kann. Diese Aussage bedarf aber noch einer genaueren Untersuchung. Wir können an dieser Stelle nur feststellen, daß die Frage nach der wirklichen Spielstärkeverteilung hier nicht beantwortet werden konnte, wichtig erscheint uns aber die Erkenntnis, daß man keinesfalls von einer ausgeglichenen Spielstärke ($s \approx 0{,}5$) der ungesetzten Teilnehmer aus der 1. Runde der Wimbledon-Turniere sprechen kann. Dies würde natürlich erst recht zutreffen, wenn man die gesetzten Spieler mit einbeziehen würde. Eine weitergehende Diskussion der Verteilungsfunktion $H(s)$ erscheint uns hier nicht sinnvoll, da im nächsten Abschnitt gezeigt wird, daß bestimmte Feinheiten der Wimbledonresultate auch mit einer beliebigen Verteilungsfunktion nicht erklärt werden können.

2.4.1 Psychologischer Faktor

Wir betrachten nun den Ablauf der ungesetzten Begegnungen der 1. Runde der Herren in Wimbledon etwas genauer, und zwar untersuchen wir, in welcher Reihenfolge die Sätze in einem Match erzielt werden. Im folgenden bedeutet die Zahl 1 "gewonnen" und die Zahl 0 "verloren". In Tab. 5 sind alle Möglichkeiten der Satzreihenfolge für ein Match mit drei Gewinnsätzen aus der Sicht des Gewinners angegeben. (Aus der Sicht des Verlierers muß 1 gegen 0 und 0 gegen 1 vertauscht werden.)

Anzahl d.Sätze	Reihenfolge	$\bar{P}_i^M$ berechnet	Wimbledon 77 - 82 Abs. Häufigkeiten 77	78	79	80	81	82	$\sum$	$\bar{P}_i^M$
3	111	$s^3+(1-s)^3$	22	22	26	20	19	23	132	45,36
4	1101	$s^3(1-s)+s(1-s)^3$	2	5	3	3	10	3	26	8,93
	1011	"	6	6	7	8	7	4	38	13,06
	0111	"	7	4	5	8	5	9	38	13,06
5	11001	$s^3(1-s)^2+s^2(1-s)^3$	3	1	0	1	1	2	8	2,75
	00111	"	0	4	2	2	2	3	13	4,47
	10101	"	4	2	2	2	0	0	10	3,44
	10011	"	2	2	1	3	4	3	15	5,15
	01101	"	0	1	1	1	0	2	5	1,72
	01011	"	2	1	0	2	1	0	6	2,06
		$\sum$	48	48	47	50	49	49	291	100

Tabelle 5
Matchhäufigkeiten. Die 1. Spalte gibt an, ob ein Match mit drei, vier oder fünf Sätzen endet. In der zweiten Spalte wird unterschieden zwischen der Reihenfolge der Sätze, 1 bedeutet "gewonnen", 0 bedeutet "verloren". Die dritte Spalte zeigt, daß die Häufigkeitsverteilung der Matchergebnisse unabhängig von der Satzreihenfolge sein sollte. In den nächsten Spalten sind die absoluten Häufigkeiten der ungesetzten Begegnungen aus der 1. Runde der Herren der Wimbledon-Turniere von 1977 bis 1982 angegeben.

Die Reihenfolge 1101 bedeutet also, daß der Gewinner den 1. und den 2. Satz gewonnen, den 3. Satz verloren und den 4. Satz gewonnen hat. Aus Fig. 4 können wir unmittelbar entnehmen, daß die Wahrscheinlichkeit, in welcher Reihenfolge ein 4-Satzmatch gewonnen wird, für alle drei Möglichkeiten exakt gleich ist, nämlich $s^3(1-s) + (1-s)^3 s^3$. Entsprechendes gilt für die sechs Möglichkeiten der 5-Satzmatche. Diese Aussage ändert sich natürlich auch dann nicht, wenn eine Spielstärkeverteilung H(s) vorliegt, da die Gleichverteilung für jedes s gilt. In Tab.5 sind die absoluten Häufigkeiten der Ergebnisse der Wimbledon-Turniere von 1977 bis 1982 eingetragen. Hier machen wir nun die überraschende Feststellung, daß die in Tab. 5 zuerst aufgeführte Reihenfolge der Viersatzergebnisse viel zu selten und die beiden anderen Reihenfolgen viel zu oft auftreten. Wir können natürlich nicht mit absoluter Sicherheit ausschließen, daß diese Abweichungen rein statistischer Natur sind. Wie unwahrscheinlich dies jedoch ist, sehen wir daran, daß in 95 % vergleichbarer Untersuchungen die Abweichungen kleiner sein müßten, wenn sie rein statistischer Natur wären. Wir haben also guten Grund zu der Annahme, daß diese Ungleichverteilung eine tiefere Ursache hat. Es wird versucht, diesen Befund mit Hilfe eines psychologischen Effekts zu erklären. Wir nehmen an, daß ein Tennisspieler, der den 1. Satz mit der Wahrscheinlichkeit s gewonnen hat, den nächsten Satz mit einer höheren Wahrscheinlichkeit $s \cdot a$ gewinnt. Hat er die beiden ersten Sätze gewonnen, so gewinnt er den dritten Satz mit der Wahrscheinlichkeit $s \cdot a^2$. a wird im folgenden "psychologischer Faktor" genannt. Das entsprechende Flußdiagramm ist in Fig. 12 angegeben. Die Zwischenergebnisse 1 : 1 und 2 : 2 werden psychologisch genauso behandelt wie der Ausgangszustand 0 : 0, ebenso werden die Zwischenergebnisse 1 : 0, 2 : 1, 0 : 1 und 1 : 2 gleichbehandelt. Aus Fig. 12 ergibt sich für die Häufigkeit der verschiedenen Matchergebnisse unter Berücksichtigung der Reihenfolge der gewonnenen und verlorenen Sätze die Tab. 6.

Sätze	Reihenfolge	$P_i^M \quad i = 3, 4, 5$
3	111	$a^3 s^3 + a^3 (1-s)^3$
4	1101	$s^3 a^2 (1-a^2 s) + (1-s)^2 \; a^2 (1-a^2 (a-s))$
	1011	$s^3 a (1-as) + (1-s)^3 \; a(1-a^2 (1-s))$
	0111	$s^2 (1-s) \; a(1-a(1-s)) + s(1-s)^2 \; a(1-as)$
5	11001	$s^3 a(1-a^2 s) \; (1-as) +$ $(1-s)^3 \; a(1-a^2 (1-s)) \; (1-a(1-s))$
	00111	$s(1-s)^2 \; a(1-a^2 (1-s)) \; (1-a \; (1-s)) +$ $(1-s) s^2 \; a(1-a^2 s) \; (1-as)$
	10101	$s^3 (1-as)^2 + (1-s)^3 \; (1-a(1-s))^2$
	10011	$s^2 (1-s) \; (1-as) \; (1-a(1-s)) +$ $(1-s)^2 s (1-a(1-s)) \; (1-as)$
	01101	$s^2 (1-s) \; (1-as) \; (1-a(1-s)) +$ $(1-s)^2 s \; (1-a(1-s)) \; (1-as)$
	01011	$s(1-s)^2 \; (1-a(1-s))^2 +$ $(1-s)^2 \; s(1-as)^2$

Tabelle 6
Die Matchhäufigkeiten P_i^M in Abhängigkeit von der Satzreihenfolge und dem psychologischen Faktor a, entnommen aus Fig. 12.

In Tab. 7 sind die mit s = 0,5 und verschiedenen Werten von a berechneten Einzelhäufigkeiten eingetragen. Die letzte Spalte enthält die Wimbledon-Resultate. Man sieht, daß der Wert a = 1,22 ausgezeichnete Übereinstimmung liefert. Der Wert a = 1,22 bedeutet also, daß ein Spieler, der den 1. Satz gewonnen hat, den 2. Satz mit 22 % höherer Wahrscheinlichkeit gewinnt usw.. Uns erscheint dieser Wert etwas hoch, deshalb diskutieren wir im nächsten Abschnitt ein Modell mit zwei verschiedenen Parametern.

Satz	a				Wimbledon 77 - 82
	1	1,15	1,22	1,25	
111	25	38	45,4	48,8	45,4
1101	12,5	11,2	9,5	8,5	8,9
1011	12,5	12,2	11,9	11,7	13
0111	12,5	12,2	11,9	11,7	13
11001	6,25	4,14	3,0	2,6	2,8
00111	6,25	4,14	3,0	2,6	4,5
10101	6,25	4,5	3,8	3,5	3,4
10011	6,25	4,5	3,8	3,5	5,15
01101	6,25	4,5	3,8	3,5	1,7
01011	6,25	4,5	3,8	3,5	2,1

Tabelle 7

Die relativen Häufigkeiten der Matchergebnisse in Abhängigkeit des psychologischen Faktors a. Die letzte Spalte enthält die Ergebnisse der ungesetzten Begegnungen der 1. Runde aus den Wimbledon-Turnieren von 1977 bis 1982. Die beste Übereinstimmung wird mit dem Wert a = 1,22 erzielt.

Wir glauben, daß es hier zum ersten Mal gelungen ist, einen psychologischen Effekt des Tennisspiels quantitativ nachzuweisen und mit Hilfe eines sehr einfachen Modells erklären zu können.

2.4.3 Modell mit zwei Parametern

Die Summenhäufigkeiten der Matchergebnisse von Wimbledon können sowohl durch eine Spielstärkeverteilung als auch durch einen psychologischen Faktor erklärt werden. Es kann auch keiner der beiden Einflüsse ausgeschlossen werden. Es ist sehr wahrscheinlich, daß ein Modell mit beiden Einflüssen die Wahrheit am besten beschreibt. Der einfachste Fall dieses Modells ist die Erweiterung des psychologischen Modells in dem Sinne, daß s nicht zu 0,5 festgesetzt wird, sondern als freier Parameter behandelt wird. Die beste Übereinstimmung erhält man mit $a = 1{,}1$ und $s = 0{,}67$. Wir glauben, daß dieser Wert von a der Wahrheit sehr nahe kommt und daß der Wert von s zumindest ein gutes Maß für den Mittelwert der Spielstärkeverteilung darstellt. Wir verzichten hier auf eine weitere Diskussion dieses Modells, da hierfür mehr statistische Daten und genauere Kenntnisse der Spielstärkeverteilung vorliegen müssen.

3. Die mittlere Zahl der Punkte pro Spiel

Wir nehmen an, daß ein Spieler mit der Wahrscheinlichkeit p einen Punkt in seinem Aufschlagspiel gewinnt. Wir können vier verschiedene Größen unterscheiden:

1. $\bar{n}_w^w$: Mittlere Anzahl der Gewinnpunkte in einem gewonnenen Aufschlagspiel
2. $\bar{n}_l^w$: Mittlere Anzahl der Verlustpunkte in einem gewonnenen Aufschlagspiel
3. $\bar{n}_w^l$: Mittlere Anzahl der Gewinnpunkte in einem verlorenen Aufschlagspiel
4. $\bar{n}_l^l$: Mittlere Anzahl der Verlustpunkte in einem verlorenen Aufschlagspiel

Der obere Index bezeichnet also die Art des Spiels, der untere Index gibt die Art der Punkte an (w: gewonnen, l: verloren).

Die allgemeine Formel für Mittelwerte lautet:

$$\bar{n} = \sum_i n_i P_i \tag{12}$$

oder speziell auf einen der oberen Fälle bezogen:

$$\bar{n}_w^l = \sum_{i=1}^{N_l} n_{w,i}^l \; P_i^l \qquad \text{mit} \sum_{i=1}^{N_l} P_i^l = 1 \tag{13}$$

Dabei bedeutet:

$n_{w,i}^l$: Gewonnene Punkte in einem Aufschlagspiel, das mit der Wahrscheinlichkeit P_i^l verloren wird.

Aus Fig. 1 können wir die oben genannten vier Größen leicht ableiten, wobei g durch Gleichung (3) gegeben ist:

$$\begin{aligned}\bar{n}_w^w &= \frac{1}{g}\left(4p^4+4\cdot 4p^4(1-p)+4\cdot 10p^4(1-p)^2+5\cdot 20p^5(1-p)^3+\ldots\right)\\ &= \frac{1}{g}\left(4p^4+16p^4(1-p)+\frac{20p^4(1-p)^2(2-3p(1-p))}{(1-2p(1-p))^2}\right)\end{aligned} \tag{14}$$

$$\begin{aligned}\bar{n}_l^w &= \frac{1}{g}\left(4p^4(1-p)+2\cdot 10p^4(1-p)^2+3\cdot 20p^5(1-p)^3\;\ldots\ldots\right)\\ &= \frac{1}{g}\left(4p^4(1-p)+\frac{20p^4(1-p)^2(1-p(1-p))}{(1-p(1-p))^2}\right)\end{aligned} \tag{15}$$

$$\begin{aligned}\bar{n}_w^l &= \frac{1}{(1-g)}\left(4p(1-p)^4+2\cdot 10p^2(1-p)^4+3\cdot 20p^3(1-p)^5\;\ldots\ldots\right)\\ &= \frac{1}{(1-g)}\left(4p(1-p)^4+\frac{20p^2(1-p)^4(1-p(1-p))}{(1-p(1-p))^2}\right)\end{aligned} \tag{16}$$

$$\bar{n}_1^1 = \frac{1}{(1-g)} \left(4(1-p)^4 + 4\cdot 4p(1-p)^4 + 4\cdot 10p^2(1-p)^4 + 5\cdot 20p^3(1-p)^5 \;....\right)$$
$$= \frac{1}{(1-g)} \left(4(1-p)^4 + 16p(1-p)^4 + \frac{20p^2(1-p)^4(2-3p(1-p))}{(1-2p(1-p))^2}\right) \qquad (17)$$

Diese vier Mittelwerte wurden in Fig. 13 abhängig von p aufgetragen. Zunächst sehen wir eine logische Symmetrie dieser Darstellung, die als Quasispiegelung um p = 0,5 ausgedrückt werden kann. Man kann sich leicht überzeugen, daß definitionsgemäß folgende Beziehungen bestehen:

$$\bar{n}_w^w\,(p) = \bar{n}_1^1\,(1-p)$$
$$(18)$$
und $$\bar{n}_w^1\,(p) = \bar{n}_1^w\,(1-p)$$

Auffallend ist, daß die Größen $\bar{n}_w^w$ und $\bar{n}_1^1$ fast gleich sind und sehr schwach von p abhängen. Für $p \rightarrow 0$ und $p \rightarrow 1$ haben sie den Grenzwert 4 , was auch ohne Rechnung sofort verständlich ist. Für p = 0,5 sind beide Größen gleich und haben den Wert 4,625, d.h. für ein gewonnenes Spiel benötigt man im Mittel nur 4,625 Gewinnpunkte, während man bei einem verlorenen Spiel natürlich genauso viele Punkte verliert.

Wir diskutieren als nächstes die Größe $\bar{n}_1^w$, das ist die mittlere Anzahl der Verlustpunkte in einem gewonnenen Spiel. Für $p \rightarrow 0$ geht diese Größe gegen den Wert 1.6. Wie kommt dieser Wert zustande? Für sehr kleines p ($<0,1$) genügt es, in der 1. Zeile der Gl. 15 nur die Terme mit p^4 zu berücksichtigen, also

$$\lim_{p \rightarrow 0} \bar{n}_1^w = \lim_{p \rightarrow 0} \frac{4p^4 + 2\cdot 10p^4(1-p)}{15\,p^4} = 1,6 \qquad (19)$$

Wir können aus dieser Gleichung sofort entnehmen, daß ein Spieler, der sehr stark unterlegen ist, von 15 gewonnenen Spielen ein Spiel ohne Gegenpunkt, 4 Spiele mit einem Gegenpunkt und 10 Spiele mit zwei Gegenpunkten gewinnt.

Es ist klar, daß ein stark unterlegener Spieler nur selten ein Spiel gewinnt, wenn er aber gewinnt, dann mit den oben angegebenen Ergebnissen. Bemerkenswert ist, daß ein Einstand sehr unwahrscheinlich wird.

Für $p \to 1$ geht $\bar{n}_l^w$ gegen 0, d.h. ein stark überlegener Spieler wird seine Spiele fast immer ohne Gegenpunkt gewinnen. Die Diskussion des Verlaufs von $\bar{n}_w^l$ ist analog. Es sei noch bemerkt, daß für $p = 0{,}5$ $n_l^w = n_w^l = 2{,}125$ ist. Das heißt, bei zwei gleich guten Spielern ohne Aufschlagvorteil ist die mittlere Anzahl der verlorenen Punkte in einem gewonnenen Spiel genauso groß wie die mittlere Anzahl der gewonnenen Punkte in einem verlorenen Spiel, nämlich gerade 2,125 Punkte (Folge obiger Symmetriebeziehung). Man kann nun leicht die Frage beantworten, wie groß im Mittel die Gesamtpunktzahl in einem Spiel ist. Wir bezeichnen diese Größe mit $\bar{n}$, sie ist gegeben durch

$$\bar{n} = g\ (\bar{n}_w^w + \bar{n}_l^w) + (1-g)\ (\bar{n}_w^l + \bar{n}_l^l) \qquad (20)$$

In Fig. 13 ist diese Größe ebenfalls eingezeichnet. Es ist klar, daß die mittlere Gesamtpunktzahl für $p \to 0$ und $p \to 1$ den Wert 4.0 anstrebt. Für $p = 0{,}5$ hat $\bar{n}$ ein Maximum mit dem Wert 6,75, $\bar{n}$ muß natürlich symmetrisch zu 0,5 sein, da es gleich ist, ob der Spieler, der mit $1 - p$ gekennzeichnet ist, oder der Spieler, der mit p gekennzeichnet ist, die Gesamtpunkt zahl ermittelt.
In Tab. 8 wurden die fünf Größen aus den beiden Wimbledon-Endspielen der Herren 1981 und 1982 für den jeweiligen Endspielteilnehmer angegeben. In Fig. 13 haben wir die Mittelwerte der fünf Größen an der Stelle $\bar{p} = 0{,}65$ (ergibt sich aus Tab. 1) eingezeichnet. Die Mittelung über alle vier Endspielteilnehmer wird gerechtfertigt durch den Umstand, daß die p - Werte für alle Teilnehmer ungefähr gleich sind. Wir können auch hier feststellen, daß die tatsächlichen Ergebnisse ausgezeichnet mit unseren Berechnungen übereinstimmen.

	$\sum_i n^w_{w,i}$	$\sum_i n^w_{l,i}$	$\sum_i n^l_{l,i}$	$\sum_i n^l_{w,i}$	N^w	N^l
Mc E	95	45	9	5	20	2
B	91	38	9	5	20	2
Mc E	91	31	20	7	21	5
C	100	44	21	11	22	4
$\sum$	377	158	59	28	83	13

$\bar{n}^w_w$	$\bar{n}^l_w$	$\bar{n}^l_l$	$\bar{n}^l_w$	$\bar{n}$
4,54	1,90	4,54	2,15	6,48

Tab. 8
Ergebnisse der Wimbledonendspiele der Herren von 1981 und 1982 (Tie-Break nicht berücksichtigt). Der obere Index bezeichnet die Art des Spiels, der untere Index gibt die Art der Punkte an (w: gewonnen, l: verloren), wobei nur die Aufschlagspiele der entsprechenden Spieler ausgewertet wurden. N^w: Anzahl der gewonnenen Aufschlagspiele, N^l: Anzahl der verlorenen Aufschlagspiele. Die Mittelwerte sind in Fig. 13 eingezeichnet. McE: Mc Enroe, B: Borg, C: Connors.

4. Zusammenfassung

Im ersten Abschnitt wurden die statistischen Grundlagen der Gewinnchance im Tennisspiel abgeleitet, wobei als wesentliches Ergebnis ein großer Verstärkungsfaktor vom Punktgewinn zum Matchgewinn gefunden wurde. Dieser besagt, daß eine kleine Überlegenheit beim Ausspielen eines Punktes zu einer großen Überlegenheit bezüglich des Matchgewinns führt. Dies ist vermutlich die Ursache für die häufig beobachtbare Fehleinschätzung der Spielstärken von Tennisspielern. Im zweiten Abschnitt wurden Häufigkeitsverteilungen von Spiel-, Satz- und Matchergebnissen berechnet und mit Wimbledon-Resultaten verglichen. Unter Berücksichtigung eines Aufschlagvorteils der Wimbledon-Teilnehmer können die Satzhäufigkeiten gut erklärt werden. Die Matchhäufigkeiten führen zu dem Schluß, daß die Spielerpaarungen in Wimbledon sehr unausgeglichen sind. Darüber hinaus sind die Matchhäufigkeiten offenbar von einem psychologischen Effekt beeinflußt, der besagt, daß ein Spieler nach gewonnenem ersten Satz den zweiten Satz mit größerer und den dritten Satz mit noch größerer Wahrscheinlichkeit gewinnt. Es wird gezeigt, daß die Matchergebnisse von Wimbledon mit einem Zweiparametermodell, das den Spielstärkeunterschied und den psychologischen Faktor enthält, völlig ausreichend beschrieben werden können. Im dritten Abschnitt wurden Mittelwerte von Punkten, die in einem Spiel erzielt werden, berechnet und mit den beiden Endspielen von Wimbledon aus den Jahren 1981 und 1982 verglichen. Auch hier hat sich ausgezeichnete Übereinstimmung ergeben. Es sei hier ausdrücklich betont, daß die einzigen Parameter, die wir für unsere Rechnungen haben, nämlich die Spielstärke p des Spielers A und die Spielstärke q des Spielers B experimentell bestimmbare Größen sind. Dies gilt natürlich nicht für den psychologischen Faktor, diesen haben wir eingeführt, ohne ihn wirklich erklären zu können.

Diese Arbeit habe ich in meiner Freizeit geschrieben. Für die Bereitschaft der Kernforschungsanlage Jülich, diese Arbeit als KFA-Report zu drucken, möchte ich sehr herzlich danken.

Den Herren R. Schätzler, Dr. H. Talarek, Dr. H. Labus und Dr. H. Grimm möchte ich für anregende Diskussionen und Frau H. Ostlender für das Schreiben des Manuskripts herzlich danken.

Meinen Kindern Monika, Carola und Michael bin ich besonders dankbar dafür, daß sie immer auf ihre Rechte verzichteten, wenn es sich bei mir um Tennis drehte.

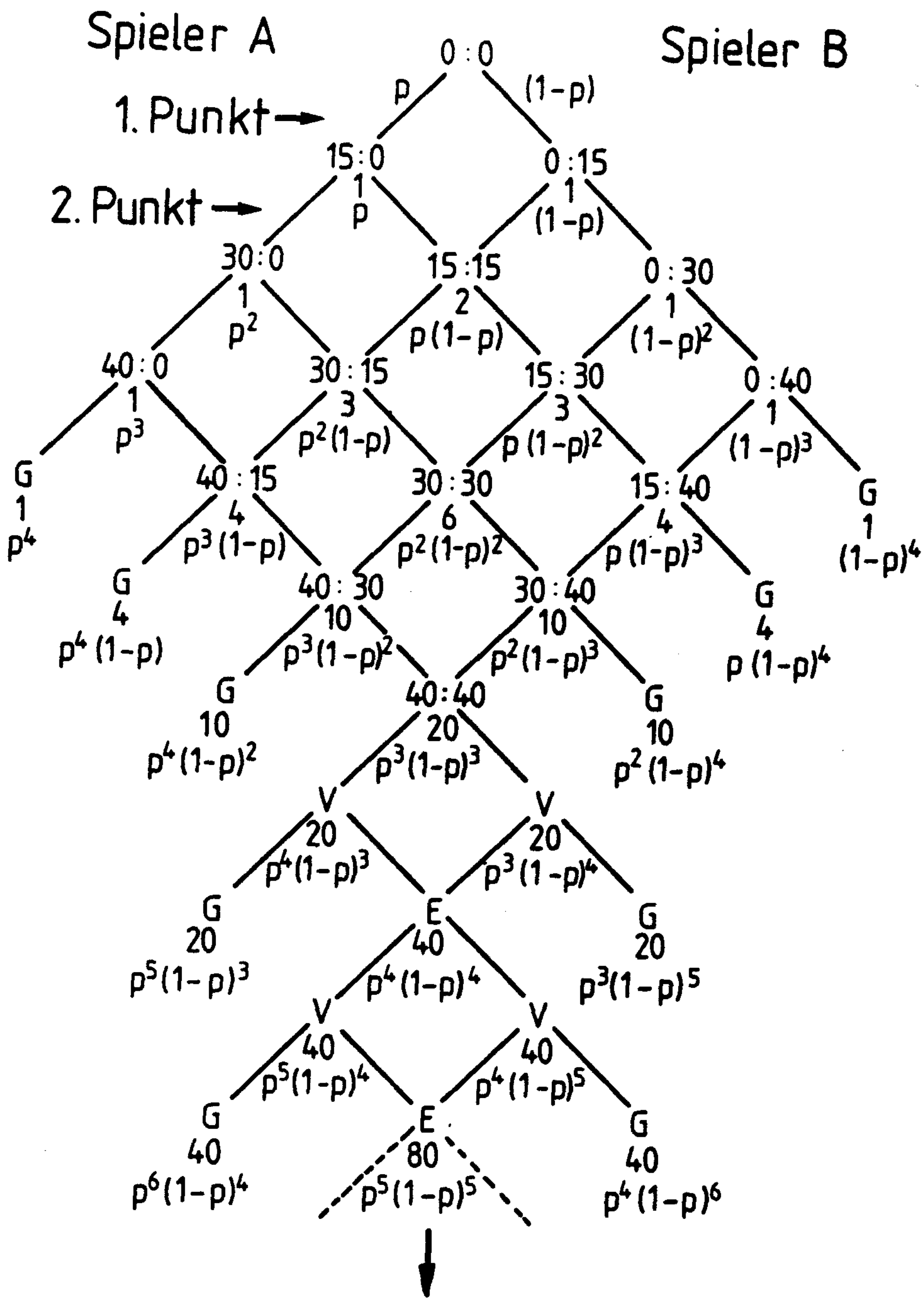


Fig. 1 - Flußdiagramm eines Tennisspiels.
Spieler A gewinnt einen Punkt mit der Wahrscheinlichkeit p, Spieler B entsprechend mit 1-p. Die Zahl unter dem jeweiligen Spielstand ist die Anzahl der verschiedenen Wege, diesen Spielstand zu erreichen. Darunter ist die Wahrscheinlichkeit angegeben, den Spielstand auf einem der möglichen Wege zu erreichen. Nach unten setzt sich das Flußdiagramm endlos fort. Bei G (game) wird das Spiel beendet. V bedeutet Vorteil, E bedeutet Einstand.

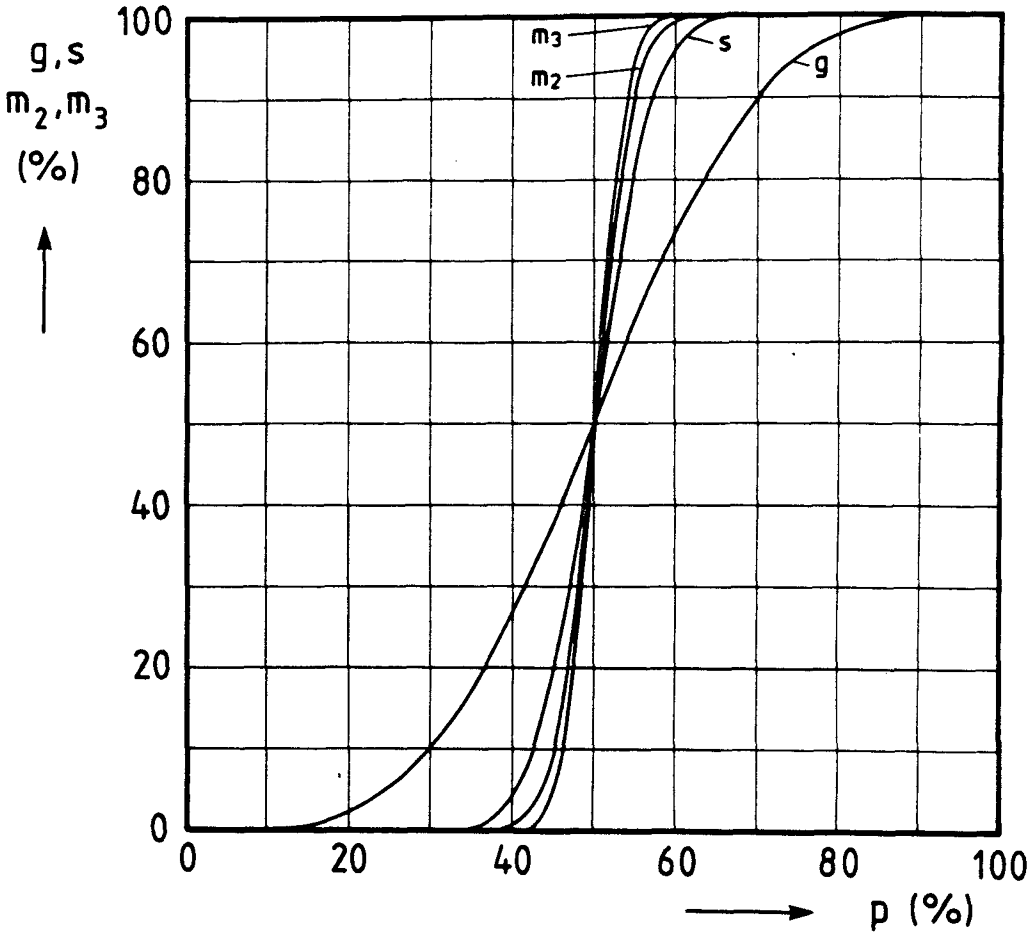


Fig. 2 - Falls ein Tennisspieler einen Punkt mit der Wahrscheinlichkeit p gewinnt, so gewinnt er ein Spiel, einen Satz, ein Match mit zwei Gewinnsätzen oder ein Match mit drei Gewinnsätzen entsprechend mit den Wahrscheinlichkeiten g, s, m_2 oder m_3.

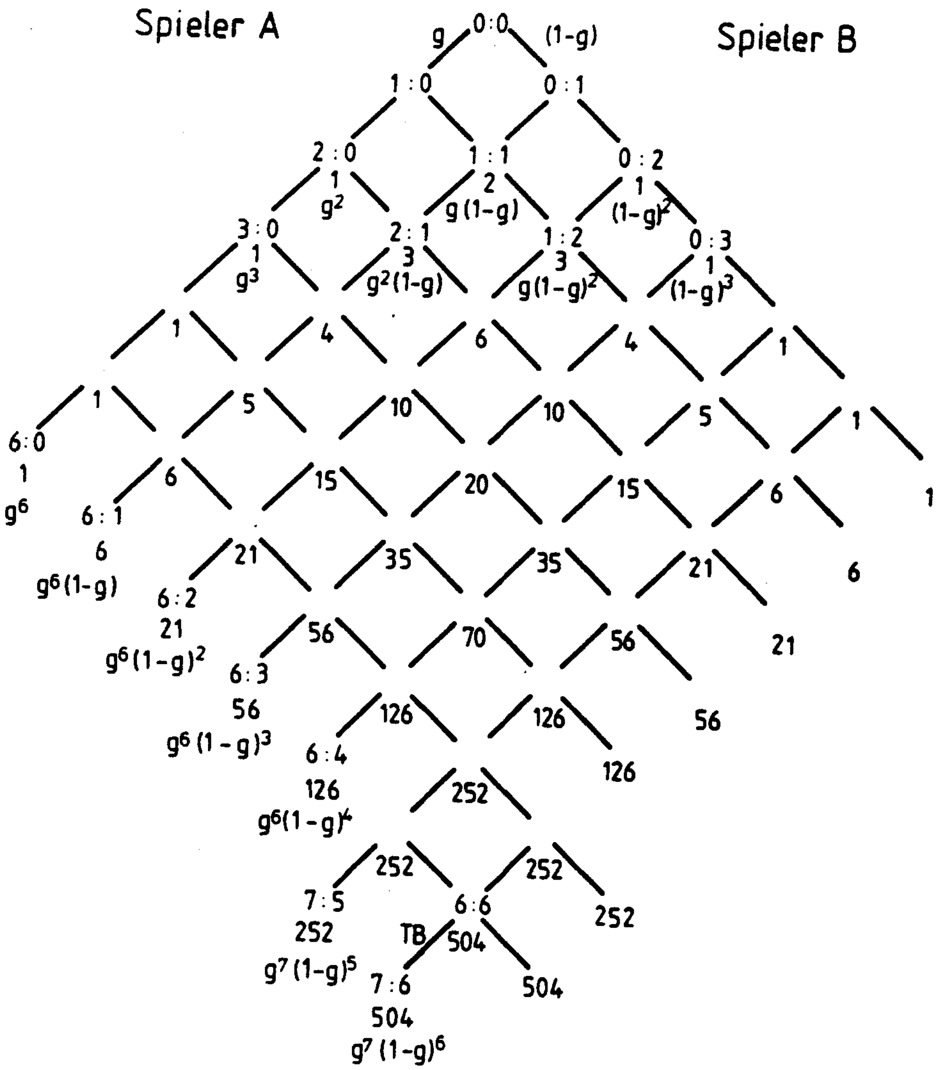


Fig. 3 - Das Flußdiagramm eines Satzes.
Spieler A gewinnt ein Spiel mit der Wahrscheinlichkeit g.

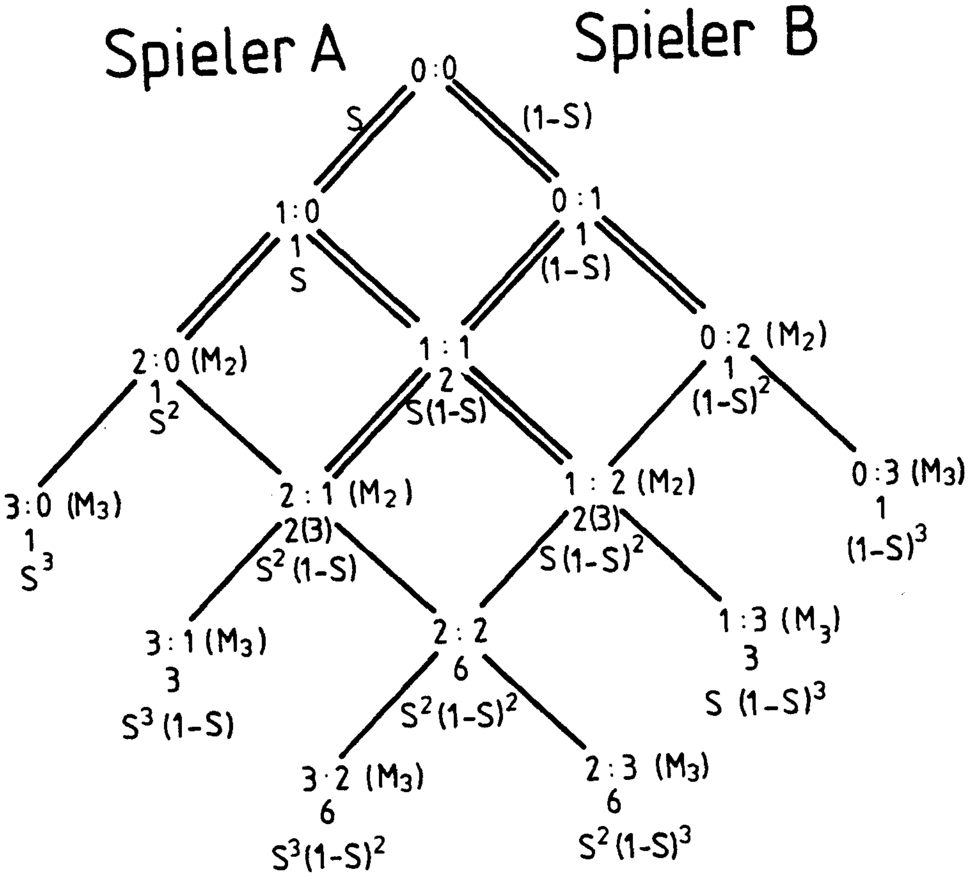


Fig. 4 - Das Flußdiagramm eines Matches mit zwei Gewinnsätzen (doppelte Linien) und mit drei Gewinnsätzen. Spieler A gewinnt einen Satz mit der Wahrscheinlichkeit s.

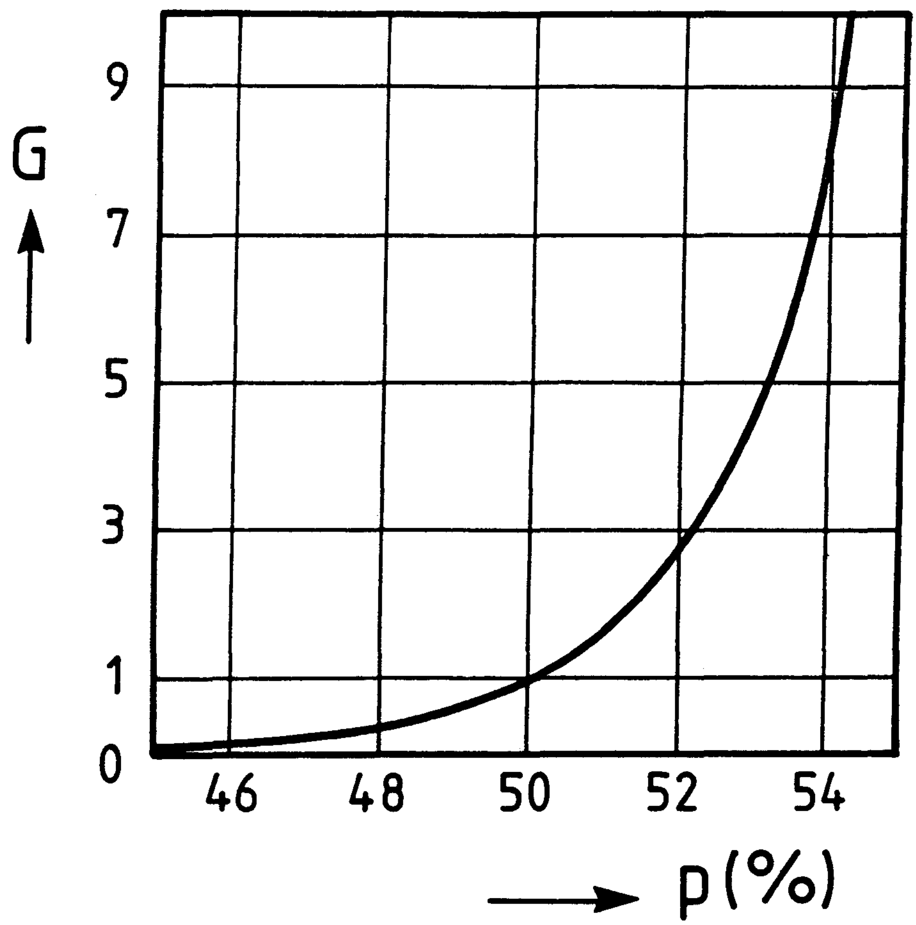


Fig. 5 - Der Verstärkungsfaktor G. Er ist definiert als das Verhältnis der gewonnenen zu den verlorenen Begegnungen, bezogen auf das Verhältnis der gewonnenen zu den verlorenen Punkten.

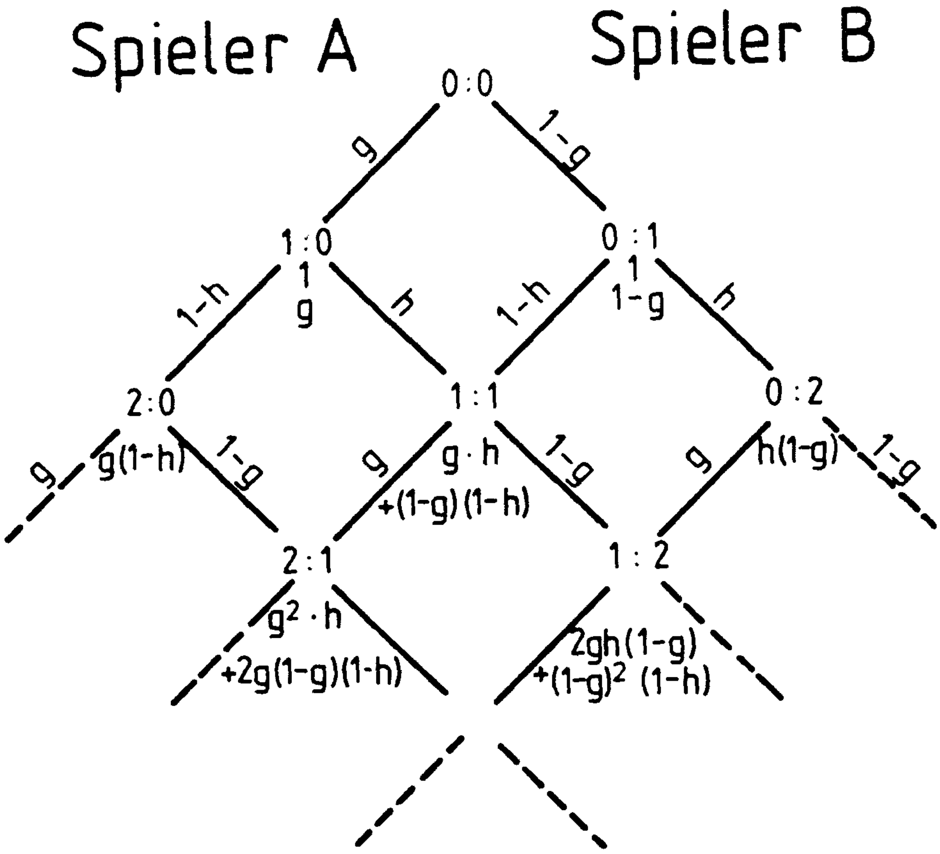


Fig. 6 - Das Flußdiagramm eines Satzes bei Berücksichtigung des Aufschlageinflusses. Spieler A beginnt mit dem Aufschlag und gewinnt jedes eigene Aufschlagspiel mit der Wahrscheinlichkeit g, Spieler B schlägt als zweiter auf und gewinnt jedes eigene Aufschlagspiel mit der Wahrscheinlichkeit h. Nach unten wurde das Diagramm abgebrochen.

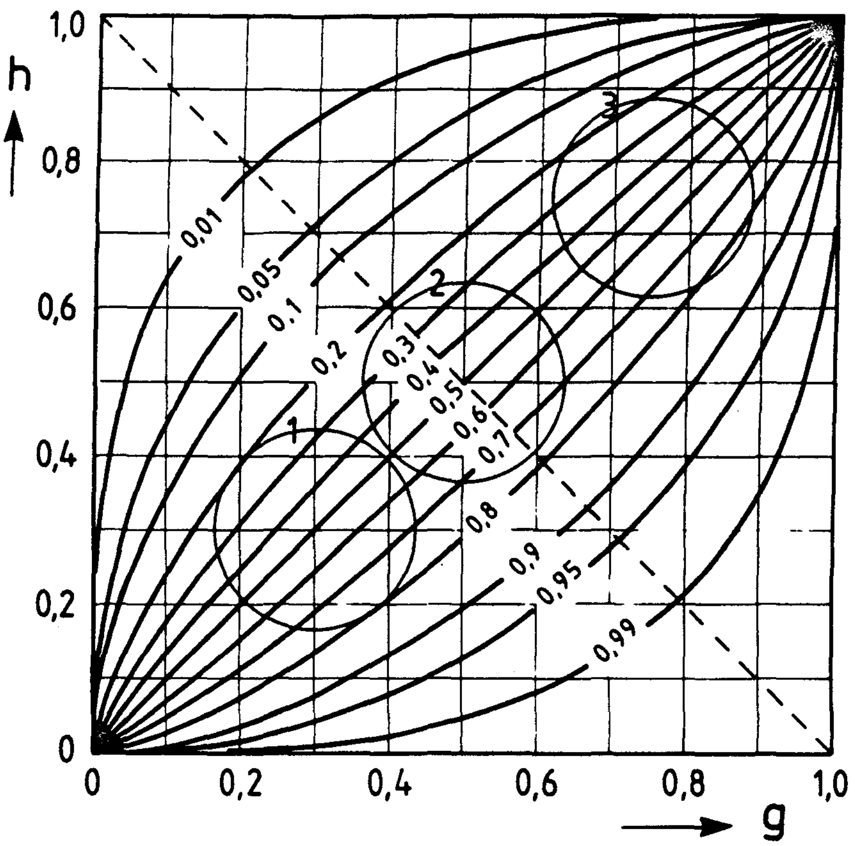


Fig. 7
Linien gleicher Wahrscheinlichkeit für Spieler A, einen Satz zu gewinnen, falls dieser sein Aufschlagspiel mit der Wahrscheinlichkeit g gewinnt und Spieler B sein Aufschlagspiel mit der Wahrscheinlichkeit h gewinnt. Die Kreise, die mit den Zahlen eins, zwei und drei gekennzeichnet sind, geben an, wo der Aufschlag der beiden Spieler von Nachteil ist, nahezu keinen Einfluß hat oder von Vorteil ist. Entsprechend sind hier die Anfänger, die mittelmäßigen und die sehr guten Tennisspieler beheimatet.

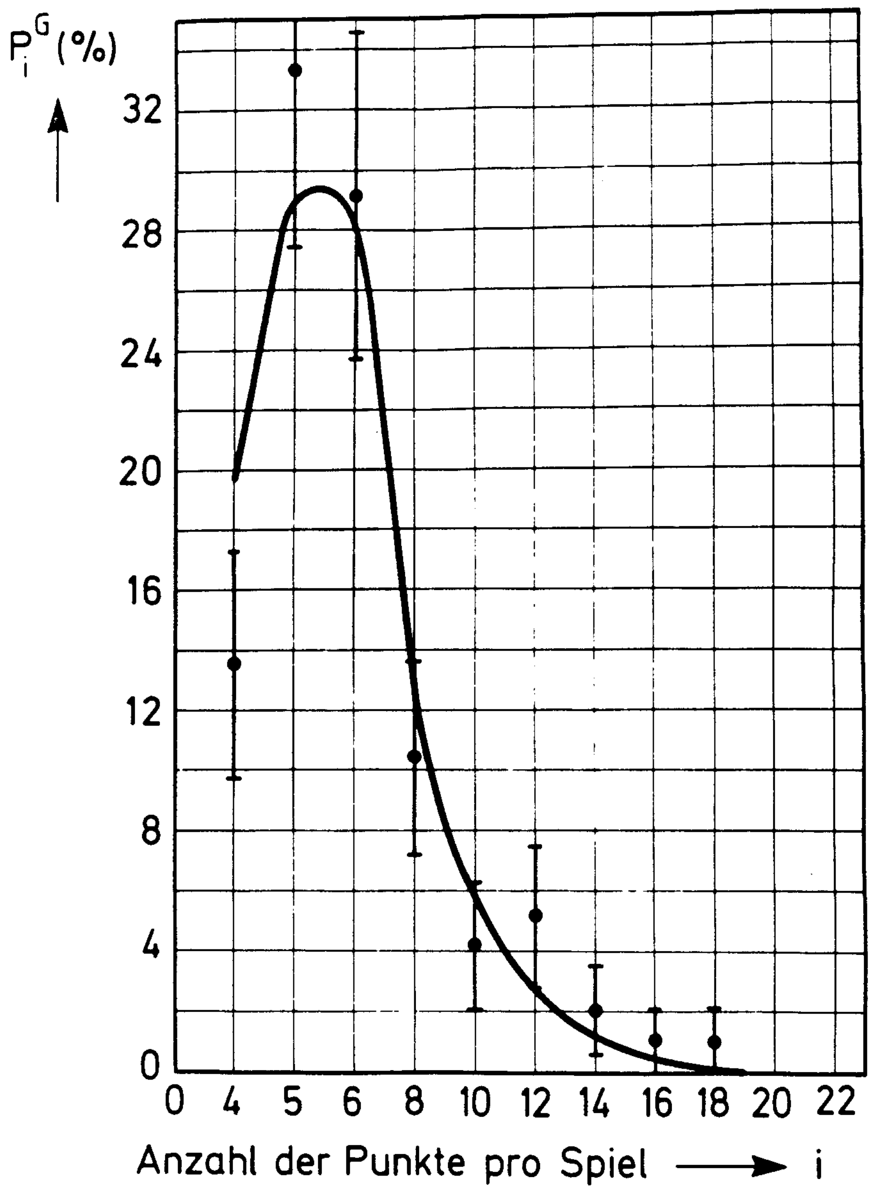


Fig. 8
Die Häufigkeit der Spielausgänge mit 4, 5, 6 ... i ... Punkten aus den Wimbledon-Endspielen der Herren von 1981 und 1982. Die durchgezogene Kurve verbindet die mit p = 0,65 berechneten Häufigkeiten.

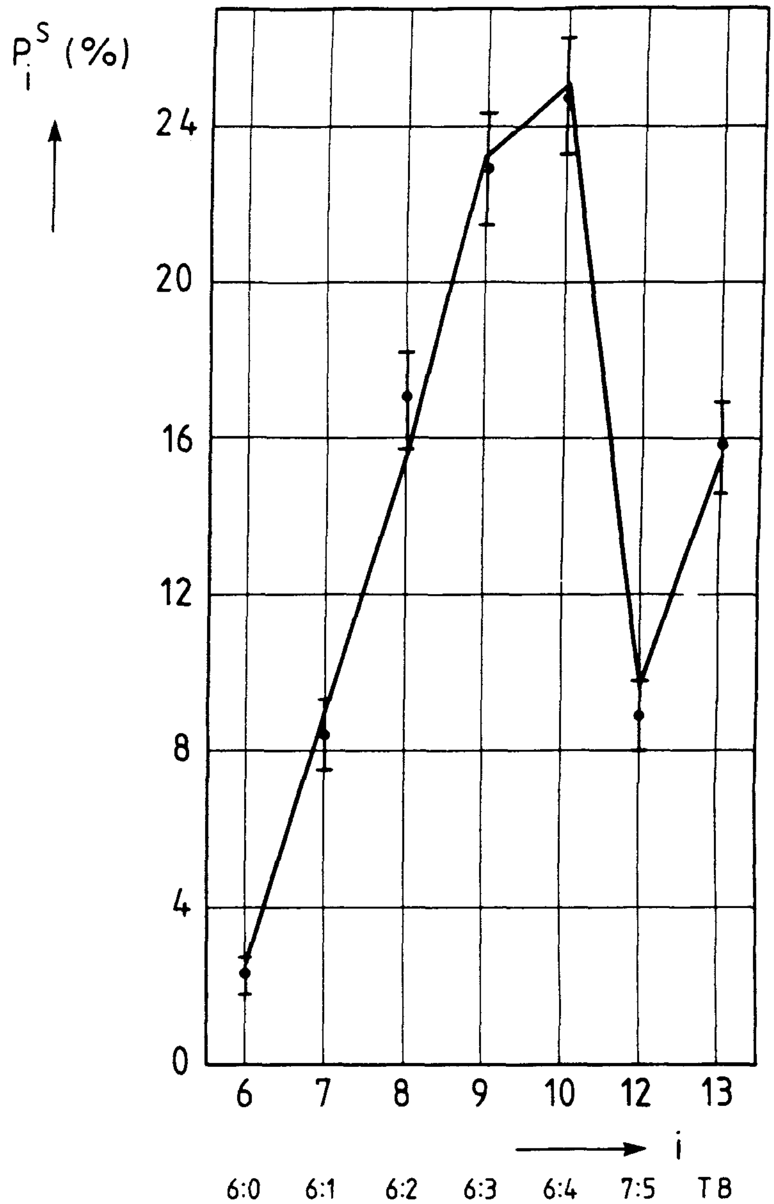


Fig. 9
Die Häufigkeiten der Satzergebnisse von allen ungesetzten Paarungen der 1. Runde der Herren aus den Wimbledon-Turnieren von 1977 bis 1982. Die durchgezogene Linie wurde berechnet.

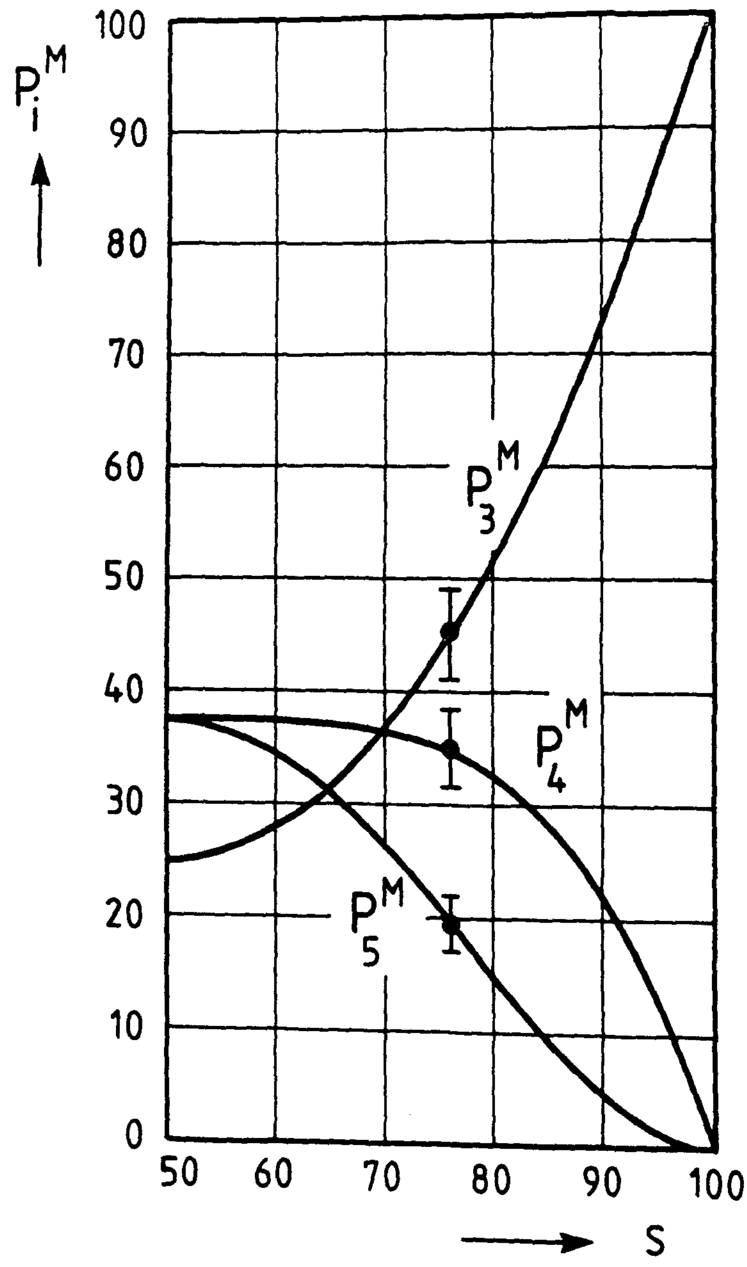


Fig. 10
Die Häufigkeitsverteilung der Begegnungen mit drei, vier und fünf Sätzen in Abhängigkeit von der Wahrscheinlichkeit s, mit der ein Spieler einen Satz gewinnt. Im Falle s = 50 (%) enden 25 % der Begegnungen mit drei Sätzen und jeweils 37,5 % der Begegnungen mit vier und mit fünf Sätzen. Die Punkte mit Fehlerbalken sind das Ergebnis der ungesetzten Paarungen aus der 1. Runde der Herren von den Wimbledon-Turnieren 1977 bis 1982. Sie werden über dem Wert s = 0,76 eingezeichnet, hier besteht die beste Übereinstimmung mit der berechneten Verteilung.

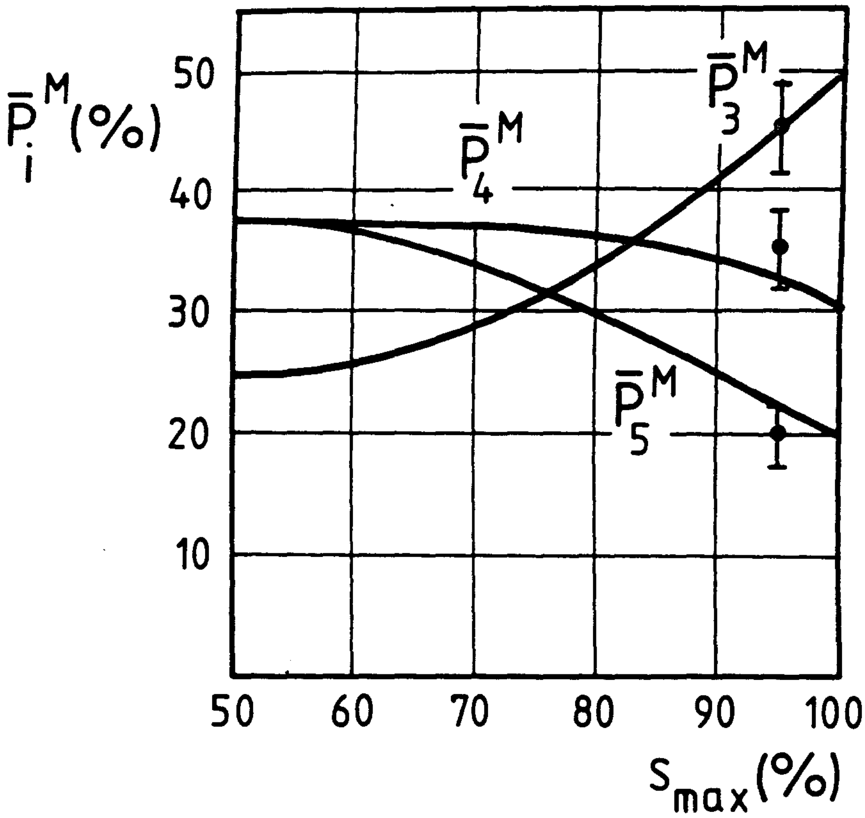


Fig. 11
Die Häufigkeitsverteilung der Matchergebnisse unter Zugrundelegung einer Spielstärkeverteilung

$H(s) = \frac{1}{s_{max}-0{,}5}$ für $0{,}5 \leq s \leq s_{max}$, sonst $H(s) = 0$

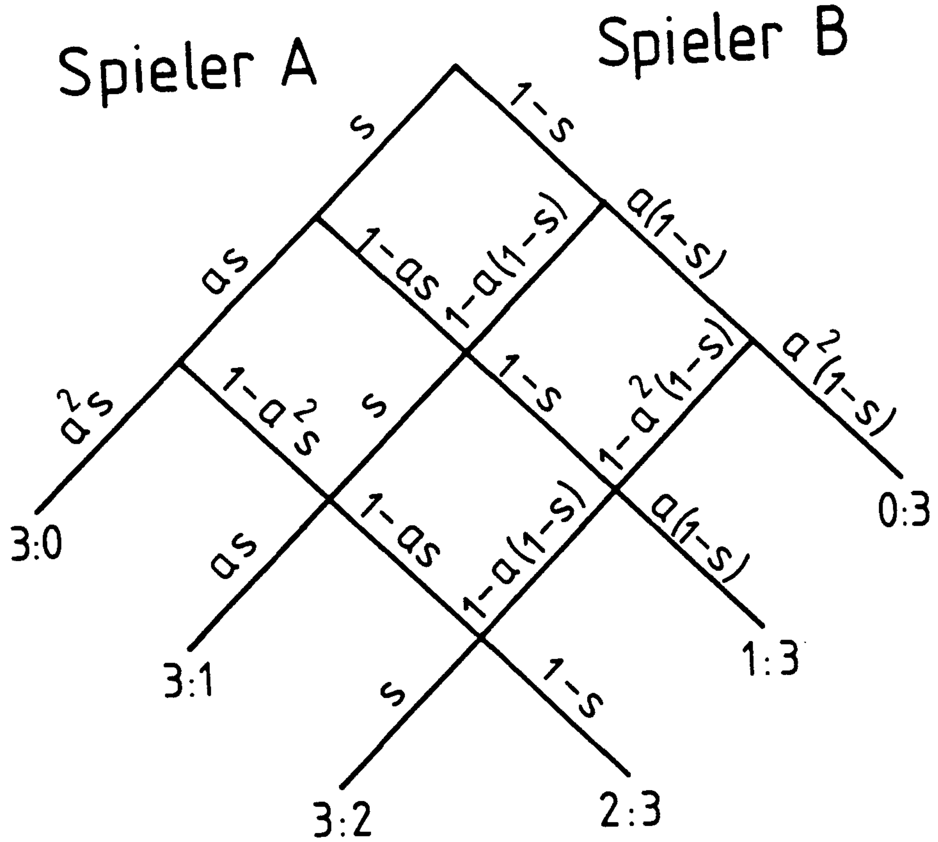


Fig. 12
Flußdiagramm eines Matches mit Berücksichtigung des psychologischen Faktors a. Wenn Spieler A den ersten Satz mit der Wahrscheinlichkeit s gewonnen hat, so gewinnt er den zweiten Satz mit der Wahrscheinlichkeit as. Wenn er auch diesen Satz gewonnen hat, so gewinnt er den dritten Satz mit der Wahrscheinlichkeit a^2s.

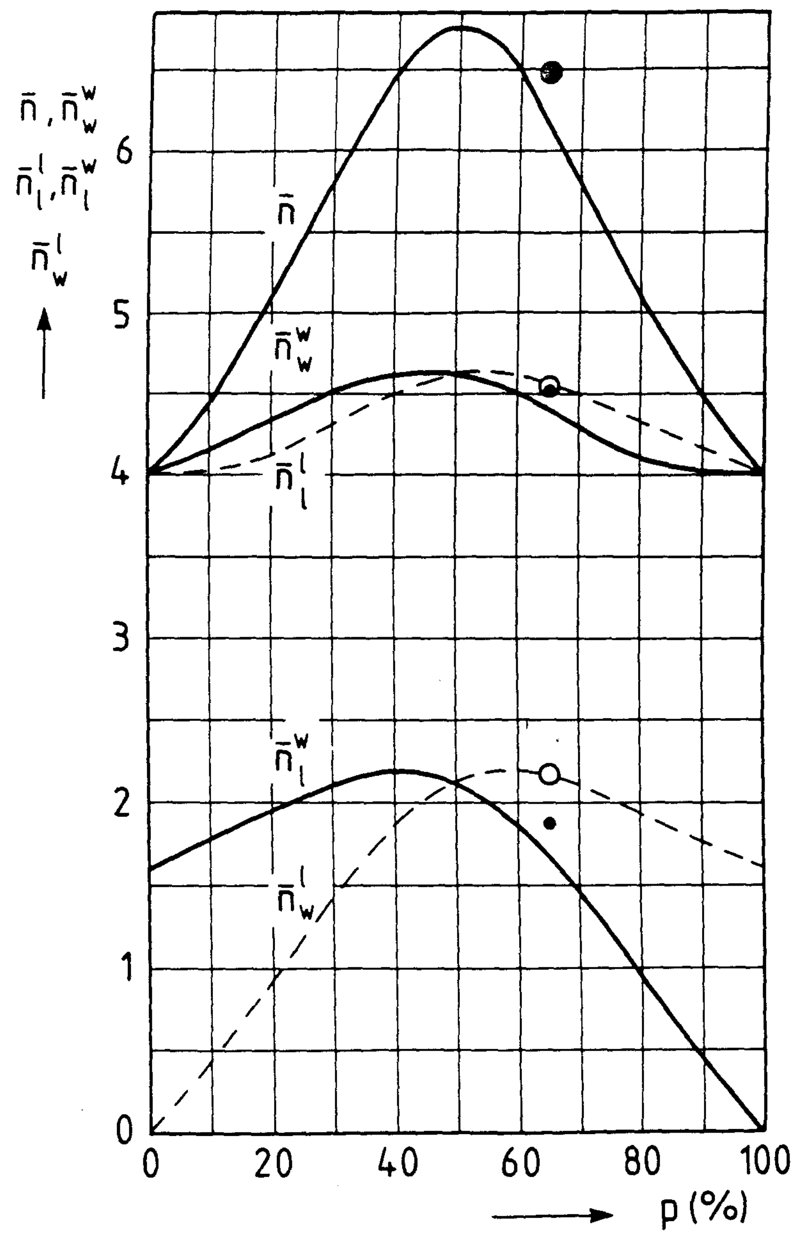


Fig. 13
Mittlere Anzahl der Gewinn- und Verlustpunkte in einem Spiel, wenn der Spieler mit der Wahrscheinlichkeit p einen Punkt gewinnt. Es bedeuten im einzelnen: $\bar{n}$: Gesamtanzahl der Punkte in einem Spiel, $\bar{n}_w^w$: mittlere Anzahl der Gewinnpunkte in einem gewonnenen Spiel. $\bar{n}_l^l$: mittlere Anzahl der Verlustpunkte in einem verlorenen Spiel. $\bar{n}_l^w$: mittlere Anzahl der Gewinnpunkte in einem verlorenen Spiel, $\bar{n}_w^l$: mittlere Anzahl der verlorenen Punkte in einem gewonnenen Spiel. Die vollen Punkte und die offenen Kreise sind aus den Wimbledon-Endspielen der Herren von 1981 und 1982 ermittelt worden. Die vollen Punkte gehören zu den durchgezogenen Kurven. Die Wimbledon-Resultate wurden an der Stelle p = 65 % eingetragen, dieser Wert wurde in Abschnitt 2.1 ermittelt.